The roles of Threonine-4 and Tyrosine-1 of the RNA Polymerase II C-Terminal Domain: New insights into transcription from *Saccharomyces cerevisiae*

Nathan Michael Yurko

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2017

Abstract

The roles of Threonine-4 and Tyrosine-1 of the RNA Polymerase II C-Terminal Domain: New insights into transcription from *Saccharomyces cerevisiae*

Nathan Michael Yurko

RNA polymerase II (RNAP II) is responsible for transcribing messenger RNAs (mRNAs) as well as non-coding RNAs such as small nuclear RNAs (snRNAs) and microRNAs in eukaryotic cells. Rpb1, the largest catalytic subunit of this complex, possesses a unique C-Terminal Domain (CTD) that consists of tandem heptad repeats (the number varying from 26 to 52 by organism) with the consensus sequence of Tyr-Ser-Pro-Thr-Ser-Pro-Ser ($Y^1S^2P^3T^4S^5P^6S^7$). The CTD is extensively phosphorylated and dephosphorylated on non-proline residues during different steps of the transcription cycle, with roles for the threonine (Thr4) and tyrosine (Tyr1) attracting more attention. For example, in chicken cells, Thr4 functions in histone mRNA 3' end formation, and Tyr1 phosphorylation is primarily associated with promoters and upstream antisense RNA formation, as well as preventing degradation of the polymerase, processes not found across all eukaryotes. A detailed introduction is described in Chapter 1.

Taking advantage of the genetic tractability of yeast cells, we created a yeast (*S. cerevisiae*) strain with all CTD threonines substituted with valines (T4V) to study the role of CTD Thr4 in transcription in yeast, which prior to this study has been poorly characterized in *S. cerevisiae*. Using the T4V strain, we found that Thr4 was required for proper transcription of phosphate-regulated (PHO) and galactose-inducible (GAL)

genes. We found genetic links between the T4V polymerase and genes encoding subunits of the Swr1 and Ino80 chromatin remodeling complexes, as well as the histone variant Htz1. We further provide evidence that CTD Thr4 is required for proper eviction of Htz1 by the Ino80 complex from genes requiring Thr4 for activation, presented in Chapter 2 of this thesis.

Finally, Chapter 3 describes the functions of CTD Tyr1 in *S. cerevisiae*. Using a strategy similar to the T4V strain, I created a strain expressing an endogenous Rpb1 with all CTD tyrosine residues mutated to phenylalanine (Y1F). We found that this strain was viable, but with a severe slow-growth phenotype. We found genetic links between the Y1F polymerase and kinase/cyclin pair Srb10/Srb11, as well as an increase in occupancy on chromatin for the same. Further analysis indicated that RNA levels of genes associated with MAP Kinase associated stressors were dysregulated, and poly(A) site selection was biased towards distal poly(A) sites. Next, using an in vitro kinase assay, we showed Tyr1 phosphorylation on the CTD by MAP kinase Slt2, and in vivo CTD Tyr1 phosphorylation levels changed based on Slt2-associated stress response, as well as a decrease in in vivo Tyr1P-RNAP II from an Slt2 kinase-dead strain. Analysis of termination factors Nrd1 and Rtt103 showed transcription termination defects were likely the result of disruption of the interaction between the CTD interacting domains of these two proteins and the Y1F CTD. Extending this, we found additional disruptions in Slt2 recruitment to chromatin, increasing the depth of our knowledge of the interplay between induction of stress-associated genes, Slt2 function, and Nrd1-mediated termination.

TABLE OF CONTENTS

Chapter 2

Chapter 3

List of Figures

Acknowledgments

I would like to begin by thanking my thesis research advisor, James Manley, for his support and leadership over the years. I have learned much from working in his lab: how to think about my experiments, how to present and defend my ideas, and how to write and succeed in the scientific field. I also thank my prior committee members Elizabeth Miller and Songtao Jia for their support, suggestions, and critical analysis of my work in the early stages. I also thank my other committee members, Lars Dietrich, Ron Prywes, Robert Fisher, and Stewart Shuman for taking the time to be on my thesis defense committee, and Lars especially for the additional help during the process.

I enjoyed my tenure in our lab with so many wonderful and exciting people. In the beginning, I worked closely with Emanuel Rosonina, who taught me how to build the mutant strains necessary for this work, and who worked with me closely to bring our ideas to fruition. Without his support, I would not be where I am today. I am thankful for the useful discussions with many lab members, especially Patricia Richard, Jian Zhang, and Tristan Coady. Whether discussing the particulars of an assay with Jian, translating my findings to other applications with Patricia, or discussing the larger ramifications with Tristan, I am happy to have made this journey with such intelligent people at hand. Other lab members have aided me in many ways – Chong Han Ng for the discussions of our work and his critical mind, Shuang Feng for her friendliness in our section, etc.

My family has supported me in this long journey towards a PhD, and my wife especially. I could not have done it without their support.

Dedication

To my family, who have given me the tools to succeed.

To my loving wife Jianyuan, who has given me the support and strength to see this to the end.

And to my daughter Chloe, who has given me the inspiration for a better future.

Preface

This thesis is divided into three chapters. Chapter 1 is an introduction to the research on the CTD of RNA polymerase II, specifically on Ser7, Thr4, and Tyr1, and parts will be published as a review article in *Transcription*. Chapter 2 is a research article entitled “Threonine-4 of the budding yeast RNAP II CTD couples transcription with Htz1-mediated chromatin remodeling”, published in *PNAS* (as Rosonina et al. 2014). Chapter 3 is a manuscript with a title “MPK1/SLT2 links multiple stress responses with gene expression in budding yeast by phosphorylating Tyr1 of the RNAP II CTD” that has been submitted for publication.

Chapter 1

The RNA polymerase II CTD "orphan" residues: Emerging insights into the functions of Tyr-1, Thr-4 and Ser-7

Nathan M. Yurko and James L. Manley

Department of Biological Sciences

Columbia University

New York, NY 10027, USA

Abstract

The C-terminal domain of the RNA polymerase II largest subunit (CTD) consists of a unique repeated heptad sequence of the consensus Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7. An important function of the CTD is to couple transcription with RNA processing reactions that occur during the initiation, elongation and termination phases of transcription. During this transcription cycle, the CTD is subject to extensive modification, primarily phosphorylation, on its non-proline residues. Reversible phosphorylation of Ser2 and Ser5 is well known to play important and general functions during transcription in all eukaryotes. More recent studies have enhanced our understanding of Tyr1, Thr4 and Ser7, and what have been previously characterized as unknown or specialized functions for these residues has changed to a more fine-detailed map of transcriptional regulation that highlights similarities as well as significant differences between organisms. Here, we review recent findings on the function and modification of these three residues, which further illustrate the importance of the CTD in precisely modulating gene expression.

Introduction

RNA polymerase II (RNAP II) is responsible for transcription of all mRNAs as well as a large and seemingly growing number of non-coding (nc) transcripts. RNAP II is a highly conserved, multiprotein complex consisting of twelve subunits, the largest of which is Rpb1. The C-terminal domain of Rpb1 (CTD) is a unique protein domain consisting of a series of tandem heptad repeats of the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser (YSPTSPS), although the exact number of repeats and deviation from this sequence varies between organisms. All steps in the synthesis of a mature

mRNA, or ncRNA, involve the CTD, as many different transcription and RNA processing factors associate with it in a dynamic manner throughout the transcription cycle. Modification of the CTD, principally phosphorylation, is important for many CTD functions. Phosphorylation occurs on all the non-proline residues, again in a dynamic manner with each residue/modification playing distinct roles. Two of these residues, Ser2 and Ser5, have garnered the most attention, and their functions are thus the best understood. Here we highlight the residues Tyr1, Thr4 and Ser7, which also play significant but perhaps more specialized roles in RNAP II transcription and regulation. A number of excellent reviews have been published concerning CTD modification and function, and the reader is referred to these for a broader picture of this unusual and still in many ways enigmatic protein domain (Hsin and Manley 2012, Corden 2013, Heidemann et al. 2013, Bentley 2014, Conaway and Conaway 2015).

The CTD has several features that allow for fine-tuned regulation of polymerase function. The length itself varies from 26 to 52 heptad repeats, with the exact number of repeats depending on the species; budding yeast has a 26 repeat CTD with little variation, while vertebrate species, including humans, have 52 repeats (Corden 1990). These longer CTDs have more divergence from the consensus sequence in their more C-terminal repeats, whereas repeats closer to the N-terminus have less variation (reviewed in Liu et al. 2010, Hsin and Manley, 2012). Additionally, a 10-residue sequence is present at the C terminus of vertebrate CTDs that helps stabilize it (Chapman et al. 2004). The actual variation in CTD composition changes between species considerably, though. While CTDs from vertebrates show considerable similarity to each other, the CTD from *D. melanogaster* displays much more variation,

with only 3 of 45 repeats matching the consensus and with considerable divergence from a related organism, *A. aegypti* (Corden 2013).

The CTD can be modified in multiple ways, more so in metazoans than in yeasts. Besides phosphorylation, the threonine and serine residues can be glycosylated (Zeidan and Hart 2010), and certain specific lysine residues can be ubiquitinated (Li et al. 2007) as well as methylated (Sims et al. 2011; Dias et al. 2015). These additional modifications occur primarily in metazoans, at least in part because *S. cerevisiae* and *S. pombe* CTDs do not have the non-consensus lysine residues. Cis/trans isomerization of the prolines functions to alter binding of proteins associated with the CTD; the peptidylprolyl-cis/trans-isomerase Ess1 in yeast (Pin1 in mammals) performs this function by binding to a Ser5P CTD, and by changing CTD conformation it can alter CTD phosphorylation dynamics (Hanes 2014).

Most early studies focused on the processes affected by Ser2 and Ser5. Two recent studies employing mass spectrometry to analyze phosphorylation seem to vindicate this approach, as these residues were shown to be the most heavily modified, at least under the conditions analyzed (Suh et al. 2016, Schuller et al. 2016). As "orphaned" residues, Tyr1, Thr4 and Ser7 have become more of a focus in recent years, with a growing body of work dedicated to determining their function. We review below studies on these three residues and their place in modulating RNAP II function. Interesting themes that emerge are that these residues and their phosphorylation seem to affect more specialized genes and processes than Ser2/Ser5, and that significant differences in the function of these residues exist between species.

Serine-7

Ser7 was the first of the “orphan” residues to receive significant attention. A schematic highlighting important functions, and illustrating some of the major differences between mammalian cells and yeast, is presented in Figure 1. Using the so-called “α-amanitin system,” in which endogenous RNAP II is inhibited by the drug so that exogenous α-amanitin-resistant mutant forms of Rpb1 can be analyzed (Gerber et al. 1995), Egloff et al. found that a human Rpb1 derivative with all Ser7 residues mutated to A (S7A) was specifically defective in expression of certain small nuclear RNAs (snRNAs), in particular the spliceosomal U1 and U2 snRNAs (Egloff et al. 2007). Additionally, Ser7 phosphorylation was shown to facilitate interaction with Integrator, a 12 subunit complex that functions in 3’ cleavage of snRNAs (Baillat et al. 2005, Egloff et al. 2007). Further studies showed that before Integrator binds, Ser7P is necessary to recruit the Ser5 phosphatase RPAP2, which stably associates with Integrator after binding the CTD (Egloff et al. 2010, 2012; Egloff 2012). Once RPAP2 is recruited, Ser5P is removed, leaving primarily a Ser7P/Ser2P CTD; Integrator subunit Int11 requires both Ser7P and Ser2P for efficient binding (Egloff et al. 2010). The complexity of the proposed “CTD code” (Buratowski 2003) increased with better understanding of Ser7, as instead of general transcription mechanisms (as impacted by Ser2 and Ser5), the idea that a CTD residue can affect a specialized aspect of transcription, i.e., U snRNA 3’ end formation, came about.

More recent studies have complicated somewhat the view of Ser7 in snRNA gene expression. Analysis of a chicken DT40 cell line expressing as the only source of Rpb1 a tetracycline (tet)-repressible derivative (tet-off) revealed that full viability in the presence of tet was restored by expression of an S7A derivative (Hsin et al. 2014).

Unexpectedly, U1 and U2 transcription, and 3' processing, were completely normal in these cells, which contrasts with the results described above (Egloff et al. 2007). While it is possible that the discrepancy reflects differences between chicken and humans, this seems unlikely due to the high conservation of all factors involved. Instead, the results of Egloff et al. may involve a synthetic effect between S7A and α-amanitin, as the inhibitor can lead to degradation of several proteins, including the elongation factor DSIF (Tsao et al. 2012), known to be important for snRNA expression (Mandal et al. 2004). However, other results of Hsin et al. (2014) do support, albeit indirectly, a role for Ser7 in snRNA 3' processing. While DT40 cells expressing an Rpb1 variant with only repeats 1-26 was entirely viable and snRNA processing unaffected, a derivative with repeats 27-52 was inviable and showed a dramatic and specific decrease in snRNA 3' processing. A parsimonious explanation for these results is that 20 of the 26 position 7 residues in the C terminal half of the CTD are not Ser. Since many of these are non-conservative changes, e.g., Lys, perhaps this is sufficient to disrupt Integrator interaction and hence U snRNA 3' processing.

The idea that CTD codes might be different for different organisms also arose from studies of Ser 7. For example, snRNA 3' ends in *S. cerevisiae* are not formed by Integrator, which does not exist in yeast, but instead by the Nrd1-Nab3-Sen1 (NNS) complex, which primarily recognizes CTD Ser5P (Vasiljeva et al. 2008), and conversely the NNS does not appear to exist in vertebrate cells (Egloff et al. 2012, O'Reilly et al. 2014). Even within yeasts, there appear to be differences in Ser7 function. In *S. pombe*, CTD Ser7 has been associated with mRNA capping (traditionally associated with Ser5), as Ser7P is one of the components that enables efficient binding of 5'-cap

methyltransferase Pcm1 to the CTD (St. Amour et al, 2012). Additionally, S7A mutation in *S. pombe* leads to reduced histone H3K4 and H3K36 trimethylation, and also exacerbates the effects of elongation factor Spt5 mutants (Mbogning et al. 2015). The same S7A mutation also led to de-repression of the *PHO1* gene, while the phospho-mimetic S7E caused hyper-repression, which points to a role for Ser7 in phosphate homeostasis in fission yeast (Schwer et al. 2015). However, in *S. cerevisiae*, no concrete function has yet been assigned to Ser7.

Despite the lack of a clear function, Ser7 is indeed phosphorylated in *S. cerevisiae*. Ser7P was found globally on transcribing RNAP II (Chapman et al. 2007), and the TFIIH kinase Kin28/Cdk7 was shown to phosphorylate Ser7 in vitro and to be required in vivo, and across organisms (Kim et al. 2009, Akhtar et al. 2009). As this kinase has already been associated with not only Ser5 phosphorylation but also mRNA capping (Komarnitsky et al 2000, Schroeder et al 2000), that similar associations with Ser7 phosphorylation exist is perhaps not surprising. Additionally, the CTD kinase Bur1 also shows Ser7 phosphorylation activity distal to promoters of highly transcribed genes, consistent with the role of Bur1 in elongation (Tietjen et al. 2010). How Ser7 is dephosphorylated is less clear. Ser7P is not a target of Rtr1 (a Ser5 phosphatase, though there is some question about this; see Xiang et al. 2012), but Ssu72 (another Ser5 phosphatase; Krishnamurthy et al. 2004) has been shown to have Ser7 dephosphorylation activity (Mosley et al. 2009, Zhang et al. 2012).

In human cells, RNAP II is known to be the target of a different modification, glycosylation (Kelly et al. 1993). The enzyme O-GlcNAc transferase (OGT) glycosylates both Ser5 and Ser7, and has been suggested to assist in proper preinitiation complex

formation (Ranuncolo et al. 2012). Despite the ubiquitous nature of Ser7 phosphorylation across eukaryotes, glycosylation does not exist in yeast, and the primary findings from work on Ser7 highlight that the variation between organisms is significant and can have important implications for the function of these residues across species. Going forward, it will therefore be wise to consider carefully differences between organisms, and how some mechanisms in one species may be repurposed or removed in another.

Threonine-4

Differences between organisms' use of the CTD became increasingly apparent as Thr4 was explored on the heels of Ser7. (Figure 2 highlights the functions of Thr4 across species.) For example, in chicken DT40 cells, using the same approach as described above for Ser7, Hsin et al. found that Thr4 was both essential for viability and necessary for efficient 3' end processing of non-polyadenylated replication-dependent histone transcripts, as processing (but not transcription) of these mRNAs, and recruitment of histone 3' processing factors, was reduced in cells expressing a T4V CTD derivative (Hsin et al. 2011). This immediately suggested a difference between yeast and vertebrates, as yeast histone mRNA 3' ends are polyadenylated like all other mRNAs (Marzluff et al. 2008). Hsin et al. also compared levels of poly-A+ transcripts as well as U1 snRNA levels and found no changes (setting Thr4 apart from Ser2 or Ser7 function), bolstering the case for specialization. Extending this to mammalian cells, ChIPseq analyses showed an increase in Thr4P signal in the 3' region of genes, and a T4A mutant showed a lethal phenotype (analogous to T4V in DT40 cells) as well as an apparent genome-wide elongation defect, as determined by elevated levels of promoter-

proximal RNAP II and lower levels at 3’ ends of genes (Hintermair et al. 2012). It should be noted that Hsin et al. also observed small increases in promoter-proximal RNAP II on several genes (Hsin et al. 2011). Whether the greater effects on elongation observed by Hintermair et al. were due to differences between chicken and human cells or to the use of the α-amanitin system as described above is not clear. Recently, a cell-cycle link was extended to Thr4, as Thr4-phosphorylated RNAP II was found to associate with centrosomes (across all cell-cycle phases save interphase) and the midbody in M phase HeLa cells (Hintermair et al. 2016).

While Thr4 is essential for viability in vertebrates, this is not the case in *S. cerevisiae*. Substitutions with Ala are well-tolerated by yeast, even when paired with an S7A derivative (Stiller et al. 2000). RNAP II with Thr4P was found globally on chromatin, suggesting a role in transcription, while distribution after the poly-A site was lower (Mayer et al. 2012). Consistent with this, in a subsequent study Thr4 phosphorylated RNAP II was detected exclusively in the chromatin fraction, and properties of a T4V derivative linked Thr4 to chromatin remodeling and histone eviction from promoters of genes required for survival under low-phosphate conditions (PHO genes) and for galactose metabolism (Rosonina et al. 2014). Normally, the repressive histone dimer H2A.Z/H2B is evicted from these promoters for such genes by chromatin remodeler INO80 during activation. In the T4V cells, this process was defective, and as a result the cells, which are fully viable in rich media, are inviable in phosphate-depleted or galactose-containing media. Schwer et al. found that expression of PHO genes was also dysregulated in S. pombe (Schwer et al. 2014). These results together are again

consistent with Thr4 being involved in relatively specific functions, which differ between yeast and mammalian cells.

More recent studies have however suggested broader roles for Thr4 in yeast, specifically in transcription termination and post-transcriptional splicing. Comparing Ser5P and Thr4P RNAP II interactomes, Harlen et al. found Thr4P devoid of spliceosomal components (associated with co-transcriptional splicing) (Harlen et al. 2016). Additional RNA analysis found that the T4V mutant partially retains introns in 90% of post-transcriptionally spliced genes, and a reanalysis of RNA-seq data of Rosonina et al. was consistent with a modest splicing defect. A proteomics analysis found that the known CTD-interacting protein Rtt103 (but not the exonuclease complex Rat1/Rai1, which function together in transcription termination) was found to associate with the Thr4P CTD. NET-seq analysis, which precisely maps the positions of elongating RNAP II, revealed an increase in RNAP II occupancy at poly(A) sites, where Rtt103 occupancy would be the highest, in a T4V mutant strain, similar to *rtt103Δ* cells. RNA-seq analysis replicated a global downstream shift of poly(A) site selection observed by Rosonina et al. (2014), consistent with a reduction of 3' end cleavage. More recent NET-seq experiments extended these findings to human cells, as Thr4P was found to correlate with termination in regions of protein-coding genes (Schlackow et al. 2017).

The termination defect itself comes about when Rtt103 cannot effectively bind the CTD (such as in the T4V mutant), which prevents effective RNAP II disengagement from chromatin and continued transcription past the poly(A) site. An additional study using a T4A mutant found that this termination defect also affects a subset of snoRNA

genes, and this was distinct from a known Ser2P requirement (Nemec et al. 2017). Additionally, NMR studies revealed that the Rtt103 CTD interacting domain (CID) binds to a Thr4P CTD as well as a Ser2P CTD, but the presence of both together inhibits binding (Jasnovidova et al. 2017). While no study suggests that Thr4 is necessary globally for transcription, there is some question as to how prevalent Thr4P is in yeast, with one estimate at only 2% of Ser2P levels (Suh et al. 2016) and another suggesting it is as abundant as Ser2P (Schuller et al. 2016).

There are currently two candidates, not mutually exclusive, for the Thr4 kinase. Initially, it was reported that CDK9 was necessary for Thr4 phosphorylation in DT40 cells, as inhibiting CDK9/P-TEFb with DRB or flavopiridol inhibited Thr4 phosphorylation (Hsin et al. 2011). Importantly, this did not reflect a requirement for Ser2P, as Thr4P was detected on an S2A derivative. Additionally, siRNA-mediated knockdown (KD) of CDK9 also decreased Thr4P in human 293 cells, and purified Cdk9 phosphorylated Thr4 in vitro (Hsin et al. 2014). A role for CDK9 in Thr4 phosphorylation is consistent with earlier studies on histone 3' end processing, since CDK9 KD resulted in increased RNAP II read-through of the natural 3' end site to a downstream polyadenylation signal, resulting in an increase in poly-A+ histone mRNA (Pirngruber et al. 2009). Other studies indicate that Polo-like kinases Plk3 and Plk1 are involved with Thr4 phosphorylation. Plk3, active throughout the cell cycle and playing additional roles in hypoxic stress response, although principally nucleolar (Zimmerman and Erickson 2007), was found to phosphorylate Thr4 in vitro and KD in HeLa cells reduced Thr4P by 50% (Hintermair et al. 2012). Plk1, an M-phase specific kinase associated with RNAP II in centrosomes and the midbody, was also found to phosphorylate Thr4 in vitro, and mutation of Thr4

(T4A) inhibited M phase progression (Hintermair et al. 2016). It is certainly possible that both Cdk9 and Plk1/3 target Thr4 under different conditions, as multiple kinases seem to target all other CTD residues. Indeed, Ser5 is also targeted in M phase, with the help of Pin1, by Cdc2/CylinB (Xu et al. 2003). The identity of the Thr4 kinase(s) in *S. cerevisiae*, and whether Cdk9 or Plk homologues are involved, is currently unknown.

Fcp1, well known to dephosphorylate Ser2 at gene 3' ends (Cho et al. 2001), also appears to be a Thr4 phosphatase. The enzyme was shown to dephosphorylate Thr4 in vitro, and Fcp1 KD in vivo increased Thr4P levels in 293 cells (Hsin et al. 2014).

Tyrosine-1

Determining the roles of Tyr1 phosphorylation has connected early successes with ongoing mystery. Much like Ser7 and Thr4, it was apparent very early on that significant differences between metazoan and yeast exist concerning Tyr1 (as illustrated in Figure 3). Tyr1 was discovered to be phosphorylated in mammalian cells relatively quickly, but by a kinase, c-Abl, that has no known yeast equivalent (Baskaran et al. 1993, 1999). As to function, Tyr1P was found many years later to be present on RNAP II at transcriptional enhancers and promoters in human cells, albeit in the antisense direction (Descostes et al. 2014). These findings are not limited to mammalian cells; upstream antisense RNAs (uaRNAs) were found to accumulate in DT40 cells when an Rpb1 containing Y1F substitutions was expressed (Hsin et al. 2014). This increase in uaRNA accumulation was not due to increased transcription (RNAP II levels were in fact decreased) and thus appears to be due to increased stability of these naturally unstable RNAs. As in human cells, Tyr1P levels on upstream regions affected by the Y1F CTD, but not on the corresponding downstream sense genes, were elevated. From these

studies, it appears that “antisense” RNAP II is excessively Tyr1 phosphorylated, and this in some way facilitates turnover of uaRNAs. Both of these studies, which employed either the α-amanitin (Descostes et al. 2014) or tet-off (Hsin et al. 2014) system, found that Y1F substitutions were lethal following inhibition of the endogenous Rpb1.

Unexpectedly, Tyr1P also plays an important role in maintaining CTD stability in vertebrate cells. A Y1F-derivative expressed in DT40 cells was shown to be present at lower levels compared to a WT derivative, and a low molecular weight Rpb1 form lacking the CTD was observed (Hsin et al. 2014). Remarkably, CTD stability (but not cell viability) could be completely restored by a single Tyr residue in the final heptad. Furthermore, in vitro assays showed that the CTD of a GST-CTD derivative could be completely degraded by purified 23S proteasomes, but a derivative Tyr1 phosphorylated by c-Abl was stable, implicating Tyr1 phosphorylation in CTD stability. Perhaps in keeping with this function Tyr1P was the only CTD phosphorylation found on cytoplasmic/nucleoplasmic (as opposed to chromatin-associated) RNAP II. Related results were observed in human cells, as a truncated Rpb1 apparently lacking the CTD was observed in cells expressing a Y1F derivative (Descostes et al. 2014).

The hunt for Tyr1 function in yeast cells has been more challenging. A truncated Y1F Rpb1 derivative (12 repeats) was shown to be lethal in *S. cerevisiae* (West and Corden 1995), but more recent experiments found that a similar truncated derivative was viable with only minor growth defects in *S. pombe* (Schwer and Shuman 2011). In contrast, in *S. cerevisiae* a full-length Y1F Rpb1 was viable, but displayed a severe slow growth phenotype (Rosonina et al. 2014, Yurko et al. 2017 [manuscript submitted]). The reason(s) for these discrepancies is unclear, but suggest differences not just between

yeast and vertebrates but also between budding and fission yeasts themselves. Also at play may be differences in the lengths/composition of the CTDs analyzed, which could lead to synthetic effects.

Several studies have begun to provide insight into the precise functions of Tyr1 and its phosphorylation in yeast. The pattern of Tyr1P globally was shown to resemble that of Ser2P, with the minor difference of a reduction in Tyr1P before the decrease in Ser2P at gene 3' ends (Mayer et al. 2012). By measuring binding affinities of several termination factor CIDs to CTD diheptad phosphopeptides, it was suggested that Tyr1P impairs termination factor recruitment to the CTD, specifically preventing Nrd1 association after promoter clearance and enabling Pcf11 and Rtt103 to bind to Ser2P-CTD after the drop in Tyr1P at 3' ends (Mayer et al. 2012). CIDs from all three of these proteins were shown to bind Ser 2P (or Ser5P, for Nrd1) CTD peptides but not Tyr1P-containing versions. Rtt103 in fact requires an intact and unmodified Tyr1 residue, since it can only bind efficiently to Ser2P-CTD peptides if a specific Asn residue can form a hydrogen bond with the hydroxyl group of Tyr1 (Lunde et al. 2010). Extending these results, Y1F mutation was found to disrupt Rtt103 binding to the CTD and impair Nrd1 recruitment to chromatin (Yurko et al. 2017, manuscript submitted). Aside from the Nrd1/NNS complex, the presence of Tyr1P during elongation and near gene 3' ends also promotes association of elongation factor Spt6, which has an SH2 domain that specifically recognizes Tyr1P and serves as a histone chaperone, with the CTD, thus linking the Tyr1 with elongation-promoting histone modifications (Burugula et al. 2014). While not all of these findings can extend to vertebrate cells, since the NNS complex

does not appear to exist outside of yeast, some overlap with Rtt103 (in vertebrates, Kub5-Hera) and Spt6 exists, though exactly how much is still in question.

The kinase(s) and phosphatase(s) responsible for modifying Tyr1 differ between yeast and vertebrate organisms. The protein kinase c-Abl was identified many years ago as a likely Tyr1 kinase in mammalian cells (Baskaran et al. 1993, 1999). c-Abl, a known nuclear tyrosine kinase, was shown to phosphorylate RNAP II Tyr1 efficiently in vitro. It contains both an SH2 domain that precedes its catalytic domain and is required for catalytic domain activation and efficient, near stoichiometric phosphorylation, and also a CTD-interacting domain specific for RNAP II. A c-Abl-related kinase, Arg, was also shown capable of phosphorylating Tyr1 in vitro (Baskaran et al. 1997). The precise roles of these kinases in vivo, and whether other Tyr1 kinases exist in mammalian cells, remains to be determined.

Determining the identity of the Tyr1 kinase(s) in yeast has been more challenging. This reflects in part the fact that that Tyr phosphorylation in budding yeast is very rare, constituting <0.1% total phosphorylation (Modesti et al. 2001; Chi et al. 2007). Furthermore, there is no yeast homolog of c-Abl, and the only known SH2 domain-containing protein is, interestingly, the Tyr1P-interacting protein Spt6 (Diebold et al. 2010). Recent work though has identified the MAP kinase Slt2/Mpk1 (Pearson et al. 2011) as a Tyr1 kinase in *S. cerevisiae* (Yurko et al. 2017 [manuscript submitted]). Slt2 phosphorylates Tyr1 in vitro, and in vivo Slt2 is necessary for full Tyr1 phosphorylation and also modulates Tyr1P levels during stress responses, which were found to rise following for example cell wall stress and heat shock. Interestingly, Slt2 also regulates other aspects of the transcription machinery, such as the Mediator

kinase/cyclin pair Srb10/11 (by phosphorylating Srb11) (reviewed in Strich and Cooper 2014), which is retained on chromatin in cells expressing the Y1F derivative (Yurko et al. 2017 [manuscript submitted]).

Glc7 appears to be the principal Tyr1 phosphatase in yeast. Glc7 was initially identified as a Ser-Thr phosphatase and plays a number of roles, for example as a subunit of the cleavage and polyadenylation factor (CPF) complex (Nedea et al. 2003) and as a required factor in the cell wall integrity pathway (Andrews and Stark 2000). Schreick et al. found that Glc7 dephosphorylates Tyr1P in vitro and, as a component of CPF, is required for Tyr1P dephosphorylation at the polyadenylation site, for recruitment of termination factors Pcf11 and Rtt103, and for termination (Schreick et al. 2014). These findings are consistent with other results that link both NNS (Nrd1) and Rtt103-dependent termination with control of Tyr1 phosphorylation (Mayer et al. 2012; Yurko et al. 2017 [manuscript submitted]). Another potential candidate for Tyr1 dephosphorylation, Ssu72, is also a subunit of the CPF complex (He et al. 2003), and has structural homology with a class of protein tyrosine phosphatases termed low molecular weight PTPs (Meinhart et al. 2003). However, Ssu72 does not show any Tyr phosphatase activity in vitro, preferring to dephosphorylate Ser5 (Xiang et al. 2012).

Modifying the CTD

Since the CTD is heavily phosphorylated, one of the most common strategies used to determine its functions has been to mutate the CTD to prevent or mimic phosphorylation. The earliest attempts to do this systematically utilized a plasmid-based system with several CTD substitution mutants (West and Corden 1995). By using a system where unmodified polymerase could be removed and replaced with a partially-

or fully-substituted synthetic variant, initial observations of phenotypes and viability were determined; initial findings indicated that Ser2 and Ser5 were necessary for viability. However, truncated CTD S2A transition mutants in fission yeast were found to be viable, and along with Tyr1 regulate the expression of iron uptake genes but otherwise show modest changes in global gene expression (Saberienfar et al. 2011, Schwer et al. 2014). Generating a full-length CTD S2A in *S. cerevisiae* and integrating it into the genome (instead of using a plasmid-based system) has also shown that this mutation is viable, but slow-growing and bearing similarity to a Ctk1 deletion strain (Cassart et al. 2012). Mutations in Ser5 so far have been shown to be unequivocally lethal, although attaching capping enzyme directly to the CTD restores viability, indicating that this lethality is a result of a failure to properly cap mRNA (Liu et al. 2008, Schwer and Shuman 2011). Interestingly, Ser7 has been consistently shown to be the most dispensable, as generating a full-length S7A mutant has shown little effect on mRNA levels in human cell lines, but are not able to efficiently process snRNAs or facilitate interaction with the Integrator complex (Egloff et al. 2007, Chapman et al. 2007).

Although S2A and S5A transitions were initially considered lethal, this may have been due to synthetic effects from the constructs used. It is necessary to note that the initial plasmid-based system was using truncated CTDs, which were also used to determine a minimum number of repeats for viability (approximately 8 consensus repeats) (West and Corden 1995). Using full-length chromosomally-integrated constructs can produce different results, since reductions in CTD length cause slower growth under normal conditions, whereas CTD constructs closer to wild-type length but with fewer consensus repeats grow much more favorably compared to shorter CTD

constructs unless stress is applied (Liu et al. 2010). A truncated CTD has been connected to increased synthetic effects with the Ras/PKA signaling pathway, as a truncated CTD and a RAS2 mutant showed similar defects upon nutrient deprivation (Howard et al. 2002). A shorter CTD has also been linked to changes in expression of Ty1 elements and Cdk8-mediated phosphorylation of the CTD (Aristizabal et al. 2015).

Applying this strategy of introducing a modified CTD has continued for Thr4 and Tyr1. Changing Thr4 to alanine has so far been uniformly lethal, excepting in fission yeast where it is still lethal on its own, but a double T4A-S7A mutant is viable (Schwer et al 2014), and in chicken cells the more conservative change to valine is also lethal (Hsin et al. 2011). A key change in Hsin et al. that enabled study of the T4V in chicken cells was the use of a tet-inducible system, which has been used previously to study other transcription factors (Wang et al. 1996; Takagaki and Manley 1998), but previously not RNAP II directly in a vertebrate system. The major finding from the paper, that Thr4 is required for recruiting 3' end processing factors to histone genes specifically, led to the idea that certain residues can affect fine-tuned and specific transcriptional programs, similar to the growth condition-specific patterns discussed above. In this case, the use of an inducible system is important to bypass the consequences of lethality of the change. Since that phenomenon is not present in all eukaryotes, investigating the effects of a T4V substitution in *S. cerevisiae* found instead a link between Thr4, the Ino80 complex, and the use of histone variant H2A.Z at a subset of promoters, linking RNAP II with chromatin remodeling at specific genes (Rosonina et al. 2014). Additional analysis using the T4V substitution revealed interactions between Rtt103 and CTD Thr4P, as well as a related post-poly-A

transcriptional defect leading to increased RNAP II density after poly-A sites and an increase in intron retention (Harlen et al. 2016). It is possible these findings extend to human cells as well, since Thr4P correlates with termination in regions of protein-coding genes, but the lethality of the T4V change in mammalian cells frustrates direct assessment (Schlackow et al. 2017).

Applying this strategy to Tyr1 has only recently been fruitful, since multiple groups have found that inserting a Y1F substitution CTD has generally resulted in lethality (West and Corden 1995, Descostes et al. 2014, Hsin et al. 2014). Much like the T4V substitution, a Y1F inducible construct in chicken cells also displays a lethal phenotype, but degradation of the mutant polymerase was also found, unlike other constructs (Hsin et al. 2014). Thus, a consensus repeat was added at the end, which still showed a phenotype but halted premature degradation of the polymerase. Only Y1F mutant CTDs in yeasts have been reported as being viable, and in the case of *S. pombe* no unusual growth defects aside from cold and iron sensitivity have been reported (Schwer et al. 2014). Recently, a full-length Y1F substitution CTD was integrated into the *S. cerevisiae* genome, and despite a severe growth defect was found to be viable, affecting the SLT2 pathway (and being a target for SLT2 phosphorylation) (Yurko et al. 2017 [manuscript submitted]). Part of this apparent disconnect may be that the Tyr1 residue is not used as much in yeasts compared to other organisms, as a CTD modified for mass spectrometry exhibited very little in vivo Tyr1 phosphorylation, despite the use of a recombinant Abl kinase to phosphorylate it in vitro (Suh et al. 2016). Another possible reason is that the Tyr1 residue may have a more specialized use in

yeasts, such as stress response (Schwer et al. 2014, Yurko et al. 2017 [manuscript submitted]), and so may be dispensable under normal conditions.

Chromatin Modification and the RNAP II CTD

Dynamic modifications to the RNAP II CTD go hand in hand with changes to the underlying chromatin structure of expressed genes. The histones that make up chromatin structure are the targets of several modifications, depending on the phospho-state of the CTD, the most prominent (excepting ubiquitination) being acetylation, phosphorylation, and methylation (reviewed in Bannister and Kouzarides 2011). Acetylation of lysines in the N-terminal tails of core histone proteins leads to the removal of positive charge that prevents folding of chromatin, leading to an open chromatin structure and enabling access of DNA for transcription (Verdin and Ott 2015). Phosphorylation is primarily linked to the DNA damage response, but H2B Tyr37 and H3 S10 phosphorylation have also been shown in yeast to either impair transcription or regulate the transcription of certain classes of genes, such as those required for sporulation or ribosomal biosynthetic genes (Rossetto et al. 2012; South et al. 2013). Methylation of histones can occur on all basic residues (arginines, lysines, and histidines), and lysine methylation can be either associated with transcriptional activation (ex. H3K4me3) or repression (ex. H3K36me2); control of gene expression this way is done either through methyltransferases such as through COMPASS/Set1 for H3K4 trimethylation and Set2 for methylation of H3K36, or by compacting chromatin (Miller et al. 2001; Butler and Dent 2012; Greer and Shi 2012). Other modifications exist, such as isomerizing the proline residues, which can help fine-tune transcriptional

response by preventing Set2-mediated lysine methylation (Nelson et al. 2006; Monneau et al. 2013).

When RNAPII begins transcription, CTD Ser5P is enriched at promoters and the 5' ends of active genes; this peak is gradually removed as transcription progresses and Ser2P levels increase (Mayer et al. 2010). Ser5 phosphorylation (along with H2B ubiquitination) is necessary for the recruitment of Set1, a histone methyltransferase that methylates the Lysine 4 on H3 (H3K4); this is done through the Paf1 complex recruiting COMPASS, which contains Set1 (Krogan et al. 2003). Set1 then works with the histone demethylase, Jhd2, to either set the conditions for transcriptional activation (H3K4me3) or repression (H3K4me2) (Ramakrishnan et al 2016). The need for differential methylation of H3K4 becomes apparent when cells need to activate or repress certain genes in response to different stress conditions. For example, resistance to Brefeldin A requires Set1 activity to activate expression, while exposure to diamide stress leads to repression of ribosomal biosynthetic genes coupled with H3S10 phosphorylation (Weiner et al. 2012, South et al. 2013). The recruitment of Set1 and resulting trimethylation of H3K4 due to CTD Ser5P enrichment can also lead to Nrd1 recruitment, which would lead to transcription termination as it applies to cryptic unstable transcripts and snoRNAs (Terzi et al. 2011). The choice between Nrd1-mediated termination and continuing transcription requires the interaction of other factors and is context dependent, one example being Slt2 which serves to recruit the Paf1 complex and blocks the Nrd1-Pol II interaction, mediated by Swi4/Swi6 protein complex (SBF) (Kim and Levin, 2011). This mechanism is especially important for the induction of stress-

responsive genes, which require immediate active transcription to enable proper response.

Once transcription is underway and the CTD transitions from primarily Ser5P to Ser2P, elongation associated factors are recruited to the RNAP II CTD and chromatin to enable efficient elongation. The kinase primarily associated with Ser2 phosphorylation is Ctk1 in budding yeast (CDK12 in mammals), although Bur1 has also been shown to phosphorylate Ser2 near the promoter (Sterner et al. 1995, Cho et al. 2001, reviewed in Hsin and Manley 2012). There is some overlap in their function; both kinases help recruit the RSC chromatin remodeler by interacting with SAGA and NuA4 histone acetyltransferases, and together recruit Spt6 to the 5' ends of genes (Spain et al. 2014, Burugula et al. 2014). Despite this, the two kinases are likely to represent different aspects of these functions, since deleting *CTK1* primarily affects CTD Ser2 phosphorylation closer to the 3' end, and Bur1 interacts directly with the CTD Ser5P followed by Ctk1 recruitment (Qiu et al. 2009). This represents a simplified order of events, as Bur1's interaction with the CTD Ser5P means it acts upstream of Ctk1 and Spt6 recruitment, both of which interact with other machinery such as the Paf1 complex in order to maintain stability throughout elongation (Dronamraju and Strahl 2014). Conversely, Fcp1 is considered to be the primary Ser2 phosphatase as it associates with RNAP II and has Ser2 dephosphorylation activity. In addition, cells lacking transcriptional activator *SUB1* show decreased association of Fcp1 to the CTD, changing the phospho-pattern of the RNAP II CTD (Cho et al. 2001, Kong et al. 2005, Calvo and Manley 2005).

As Ser2 is phosphorylated, the histone methyltransferase Set2 is able to directly recognize and bind to CTD Ser2P, providing the sole methyltransferase activity for H3K36 in *S. cerevisiae* (Wagner and Carpenter 2012, Tanny 2014). Besides Set2, H3K36 trimethylation also requires Spt6, as Spt6 influences Set2 protein levels and is necessary for H3K36 trimethylation when Set2 levels are normal. However, the SH2 domain of Spt6 is not required for this (Youdell et al. 2008). The chromatin modification aspect of Spt6 has also been observed in *S. pombe* since it is required for both H3K36me3 and H3K4me3 levels to repress antisense transcription (DeGennaro et al. 2013). It should be noted that Spt6, while identified initially as interacting with CTD Ser2P (see Youdell et al. 2008), has been identified as better interacting with CTD Tyr1P in yeast, consistent with the characterization of the SH2 domain and crystal structure studies (Liu et al. 2011, Close et al. 2011). Part of the confusion in *S. cerevisiae* may result from the fact that Tyr1 and Ser2 phosphorylation of the CTD have similar patterns until the 3' ends of genes, where Spt6 levels and CTD Tyr1P levels drop, and that tyrosine phosphorylation in yeast is considered rare and poorly characterized. In addition, *in vitro* binding of Spt6 to CTD Ser2P makes it harder to dissect out the specific roles of Tyr1P and Ser2P (Sun et al. 2010; Mayer et al. 2011, 2012). Despite this, the roles of Set2 and Spt6 in modifying histones to enable transcription are clear, and Ser2P on the RNAP II CTD Ser2P along with the rest of the elongating polymerase provides a point from which many factors combine and enact changes to the chromatin.

In addition to the interplay of histone and CTD modifications, there exist also some histone variants in yeast. Besides the core histones (H2A and B, H3, H4), yeast

also has an essential centromeric H3 (CSE4), nonessential H2A variant H2A.Z, and an H1 (reviewed in Eriksson et al. 2012). CSE4 (or CENP-A in mammals) is required for formation of the centromere with H4 on chromosomes, and assembly of the kinetochore, but is otherwise not involved in transcription (Quenet and Dalal 2012). H1 in yeast, while not required for viability, may be responsible for more subtle aspects of chromatin structure, such as longer chromatin loops without this histone (Georgieva et al 2012). It has also been linked to meiotic recombination during sporulation in yeast through binding to meiotic gene promoters by Ume6 (as part of the Rpd3L histone deacetylase complex), and inhibits homologous recombination in vegetative cells, although this view has been challenged recently as sporulation and meiosis were not impaired in *H1* deletion cells (Bryant et al. 2012, Brush 2015). The histone variant H2A.Z, however, is much more relevant to transcription.

Conserved from yeast to human, H2A.Z is a histone variant that replaces histone H2A in primarily promoter-associated nucleosomes (reviewed in Henikoff and Smith 2015). By placing H2A.Z so close to the transcription start site under normal conditions, the expression of genes poised for activation can be regulated at the chromatin level. H2A.Z's deposition at promoters requires Swr1 (and in some cases Tup1), whereas its removal is thought to involve Ino80, but possibly not under all conditions (Zhang et al. 2005, Gligoris et al. 2007, Tramantano et al. 2016). However, under conditions of phosphate depletion or galactose induction, H2A.Z is rapidly removed from promoters to enable transcription of target genes such as *PHO5* and *GAL1*, and the Ino80 complex in combination with Thr4 phosphorylation of the RNAP II CTD connects this chromatin remodeling with the activity of the polymerase (Rosonina et al. 2014). Besides this

activity for active transcription, H2A.Z has been suggested to affect post-transcriptional mechanisms as well; in Rrp6 mutants, H2A.Z is also found in 3' ends of genes associated with antisense transcription start sites, and H2A.Z deletion corresponds to down-regulation of these antisense transcripts (Gu et al. 2015). Reconciling these findings with each other, it is possible that Ino80's role in H2A.Z removal is limited to specific promoters, as previous experiments with upstream activation sequences of many *PHO* and *GAL* genes have shown (Ebbert et al. 1999, Barbaric et al. 2007).

Termination and the RNAP II CTD

Once RNAP II reaches the 3' end of a gene, and the CTD is primarily Ser2P, the 3' end processing machinery arrives to the 3' UTR and polyadenylates the nascent mRNA. To this end, many 3' end factors have CTD Interacting Domains (CIDs) that recognize a primarily CTD Ser2P (save for Nrd1) (reviewed in Mischo and Proudfoot 2013, Porrua and Libri 2015, Arndt and Reines 2015), but a recent addition to our understanding of this process is the role of Tyr1 of the CTD. Tyr1 phosphorylation is mostly similar to Ser2P in pattern when looking at gene-averaged ChIP profiles (Mayer et al. 2012), the difference being that Tyr1 phosphorylation decreases upstream of the pA site whereas Ser2 phosphorylation persists past the pA site. The phosphatase responsible for this change, Glc7, is part of the cleavage and polyadenylation factor (CPF) complex and acts at the polyadenylation site to cleave the mRNA and add a poly-A tail post-dephosphorylation (Schreick et al. 2014). Another phosphatase, Rtr1, has also been described as having this activity, dephosphorylating both Ser5 and Tyr1 in vitro and providing a level of autoregulation at the mRNA level through REX exonuclease activity (Hsu et al. 2014, Hodko et al. 2016). The kinase responsible for

Tyr1P to begin with is not known in yeast, but in mammalian cells c-Abl has been identified early (Baskaran et al. 1993).

With Tyr1 phosphorylation identified as modulating termination, it was found that termination factors Nrd1, Rtt103, and Pcf11 were primarily affected. The two factors Nrd1 and Rtt103 are quite different in form and function. Generally speaking, Nrd1 associates with primarily CTD Ser5P, whereas Rtt103 associates primarily with CTD Ser2P, though this is not the only factor affecting termination pathway (Vasiljeva et al. 2008, Gudipati et al. 2008, Heo et al. 2013). The termination pathway chosen between these two factors involves Pcf11, which also has a CID that favors Ser2P but is stimulated once Ser5P is dephosphorylated (Lunde et al. 2010, Grzechnik et al. 2015). However, this is a simplified view of how these determine the fate of termination pathway, as CTD phosphorylation is dynamic and these termination factors bind to CTDs with a mix of Ser5 and Ser2 phosphorylation (in the case of Pcf11) or CTDs with all serine residues phosphorylated (in the case of Nrd1) (Barilla et al. 2001, Kubicek et al. 2012). In fact, Nrd1-dependent terminators appear in regions that overlap CTD Ser5 and Ser2 phosphorylation, implying that several steps beyond detection of the CTD phosphostate are needed (Mayer et al. 2010, Tietjen et al. 2010).

Nrd1 functions as part of the Nrd1-Nab3-Sen1 complex, composed of two RNA-binding proteins (Nrd1 and Nab3) and a DNA/RNA helicase (Sen1) (Steinmetz et al. 2001, Finkel et al. 2010, Jamonnak et al. 2011). Nrd1 and Nab3 recognize specific RNA sequences through their RNA recognition motif domains (RRMs), GUA[A/G] for Nrd1 and UCUU for Nab3, which serve as terminators for snRNA and snoRNA transcripts (Steinmetz et al. 2001, Maris et al. 2005, Carroll et al. 2007). Sen1 then binds directly to

Nab3, which then enables NNS to bind to the exosome, a collection of RNases which degrades the nascent RNA (Conrad et al. 2000, Houseley et al. 2006). As the presence of the Nrd1 and Nab3 binding sites are found more often in AU-rich sequences or at sites with relatively short transcript length, it makes sense that the Nrd1 CID would be optimized for Ser5P binding (Kim et al. 2011, Porrua et al. 2012). The resulting transcripts, primarily cryptic unstable transcripts (CUTs) besides snRNAs and snoRNAs, require Sen1 to release the polymerase early (Porrua and Libri 2013) and the exosome to be recruited through the TRAMP complex (which additionally trims poly(A) tails of RNAs destined for the exosome), composed of Trf4, Air2, and Mtr4, which is done through the Nrd1 CID and Trf4's Nrd1 interacting motif (which mimics CTD Ser5P) (Vasiljeva and Buratowski 2006, Tudek et al. 2014). Of note is that the relative number of molecules of Sen1 in the cell is extremely low (~125) compared to Nrd1 and Nab3, which may indicate that Sen1 might be a limiting factor in the process (Ghaemmaghami et al. 2003).

While processing these non-coding RNAs requires the above complexes in yeast (exosome, TRAMP, and NNS), it is not clear that the homologs in human cells have all the same functions. The Sen1 homolog, Senataxin, has been implicated in neurodegenerative disorders such as ataxia and ALS (Moreira et al. 2004; Chen et al. 2004) and also resolves RNA/DNA hybrids while also associating with the exosome as in yeast (Skourti-Stathaki et al. 2011; Richard et al. 2013). No known homologs of Nrd1 or Nab3 exist in humans, though some human genes such as *SCAF8* which encodes for a RNA-binding protein, have homology with Nrd1 while not showing function of processing snRNA genes (Yuryev et al. 1996, O'Reilly et al. 2014). In contrast,

mammalian snoRNA genes are found within introns, and a different complex not found in yeast called Integrator associates with the RNAP II CTD and enables transcription of snoRNA, similar to standard mRNA processing (Baillat et al. 2005, Richard and Kiss 2006). A homolog of Nrd1, Seb1, has recently been found in *S. pombe*, however it does not function in NNS termination but instead associates with other 3' processing factors and is involved in alternative poly-A site selection, as well as ncRNA binding in the process of heterochromatin reassembly (Marina et al. 2013, Lemay et al. 2016).

Another alternative to NNS-mediated transcription termination is Poly-A Site (PAS)-dependent termination, which utilizes the cleavage and polyadenylation complex (CPAC) and is primarily associated with most protein-coding genes (reviewed in Mischo and Proudfoot 2013, Porrua and Libri 2015). Central to this is Pcf11, a CPAC component, which also determines the switch between NNS and PAS-dependent termination and is recruited both to 3' ends of genes (as part of cleavage and polyadenylation factor IA) and to Nrd1-dependent terminators (Amrani et al. 1997, Grzechnik et al. 2015). The role of Pcf11 is more prominent in PAS-dependent termination, however; the N-terminal CID interacts preferentially with the CTD Ser2P (Noble et al. 2005, Lunde et al. 2010), followed by interaction with protein dimer Rna14-Rna15 (binding to RNA) (Amrani et al. 1997, Noble et al. 2004, Paulson and Tong 2012), and then two zinc finger domains mediate the interaction with Clp1 (Ghazy et al. 2012, Yang et al. 2017). Together, these four components form cleavage and polyadenylation factor IA, and adding Hrp1, an hnRNPA/B family member which binds to RNA poly-A sequences, forms a complete cleavage and polyadenylation factor I (CFI) complex (Henry et al. 1996, Gaillard and Aguillera 2014). The RNA undergoes

endonucleolytic cleavage at the poly-A site, and a poly-A tail is added by poly-A polymerase Pap1; the Rat1 exonuclease completes the process by targeting the 5’ end of the transcript (see discussion on Rtt103) (Lingner et al. 1991, Brodsky and Silver 2000, reviewed in Porrua and Libri 2015).

Mutations in Pcf11 tend not to be tolerated by *S. cerevisiae*. The gene is essential, and mutations in either the CID or the domain that interacts with Clp1 are lethal (Sadowski et al. 2003). Specifically, mutants in the Clp1 interaction domain lead to defects in 3’ end processing and causes transcriptional readthrough (Noble et al. 2007, Ghazy et al. 2012). However, should the interaction between Pcf11 and the RNAP II CTD be unimpeded, the conformational change of the CTD is transduced through to the CFI complex and leads to its dissociation; a mutation in the CID will lead to a failure to dismantle this complex (Zhang et al. 2005). Although the human version of Pcf11 primarily shares homology at the CID, and mutations in human Pcf11 also share defects in cleavage and termination (indicating functional conservation), human Pcf11 is otherwise twice as large and less well characterized (West and Proudfoot 2008).

In contrast to Nrd1, Rtt103 is much more conserved functionally from yeast to human. For it to function, the proteins Rat1 (a 5’ RNA exonuclease) and Rai1 (decapping endonuclease) form a complex which then interacts with Rtt103, which has a CID that preferentially recognizes CTD Ser2P (Kim et al. 2004). The function of these three together form the basis for the “torpedo” model of termination, where the nuclease first processes the 5’-end of RNA after poly-A site-dependent cleavage (see section on Pcf11), and catches up with RNAP II before dissociating it from DNA (Connelly and

Manley 1988, West et al. 2004, Pearson and Moore 2013). Although it has been shown *in vitro* that Rat1 and Rai1 cannot elicit termination on their own (Dengl and Cramer 2009), a more recent study has shown that stalled elongation complexes can in fact be dissociated using these two proteins, and that Rtt103 can recruit Rat1/Rai1 protein complex to the 3' ends of genes (Lunde et al. 2010, Pearson and Moore 2013). Recent studies of the RNAP II CTD have confirmed that Rtt103 is not only preferentially associated with the Ser2P form (Suh et al. 2016), but also that Rtt103 is blocked by Tyr1 phosphorylation before the poly-A site (Mayer et al. 2012) and interacts with CTD Thr4P, possibly in the context of splicing (Suh et al. 2016, Harlen et al. 2016). Further connecting to Tyr1 phosphorylation of the CTD, it has been shown that Glc7, identified as a tyrosine phosphatase, is recruited to the 3' ends of genes and enables the recruitment of Pcf11 and Rtt103, coupling Glc7 activity with the cleavage and polyadenylation factor machinery (Schreieck et al. 2014).

The human homolog of Rat1 is Xrn2; it is also a 5'-3' exoribonuclease and is associated with various cancers, most notably lung (Lu et al. 2010, Morales et al. 2014). Also, like Senataxin, the human homolog Xrn2 and the Rtt103 homolog Kub5-Hera have been shown to be involved in DNA damage response and double strand break repair, and specifically suppresses the formation of R-loops, which are persistent RNA:DNA hybrids that occur during transcription and are parts of the termination process (Skourti-Stathaki et al. 2011, Ni et al. 2011, Morales et al. 2014). The role of Rtt103 in DNA damage response is not clear, although it is known to be involved as deletions sensitize yeast to DNA damage, specifically double strand breaks, and is implicated in a DNA damage response involving RNAP II elongation (Srividya et al.

2012, Windsor et al. 2013). One possibility is that the ability of termination factors like Rtt103 to pause and disengage the polymerase is taken advantage of in this situation, which would enable the DNA damage response to work effectively (Srividya et al. 2012). It is not surprising, then, that the phosphostate of the RNAP II CTD would be important if this was the case, and is in fact a plausible idea, since the Ser2 kinase Ctk1 is necessary for DNA damage response in this exact case, with DNA damage leading to an increase in Ser2P globally (Ostapenko and Solomon, 2003).

Concluding remarks

With the advent of new tools and strategies to explore the different functions of the CTD, more of the gaps in the knowledge surrounding how the CTD is modified over the course of transcription and how it functions are being filled in. But differences between metazoans, or even between similar species such as budding and fission yeasts, have become apparent and raise more questions. For example, what is the significance of the involvement of the INO80 complex/Htz1 and Rtt103 for Thr4 phosphorylation beyond budding yeast? The INO80 complex is evolutionarily conserved (Chen et al. 2011) and the Rtt103 homolog Kub5-Hera is well-studied (Morales et al. 2014), but how their functions impact RNAP II activity vary, and not all of these functions are conserved between organisms. A stronger example of this divergence is evident with Tyr1. Beyond affecting termination factor recruitment, what aspects of Tyr1 phosphorylation and Slt2 function are transferable to mammalian cells? Evidence that Slt2 is present with RNAP II on transcribed genes and homologs Erk1/2 are recruited to genes in a similar manner to Slt2 (Kim and Levin 2011, Mikula et al. 2016) has been reported, but might one of these kinases, or a related one, have Tyr1P activity? Future

studies will continue to describe the many intricacies of the CTD and discover new aspects of regulation, but now that the “orphans” are orphans no longer, continued comparison across organisms will prove enlightening.

Acknowledgments

We thank Chong Han Ng for critical reading of this manuscript. Work from the authors’ lab is supported by NIH grant R35 GM118136 to JLM, and NMY was partially supported by NIH training grant 5T32GM008798.

References

Adkins MW, Tyler JK. 2006. Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. *Mol Cell* **21**: 405-16.

Akhtar MS, Heidemann M, Tietjen J, Zhang D, Chapman RD, Eick D, Ansari AZ. 2009. TFIIH kinase places bivalent marks on the carboxyl-terminal domain of RNA polymerase II. *Mol Cell* **34**: 387–393.

Amrani N, Minet M, Wyers F, Dufour ME, Aggerbeck LP, Lacroute F. 1997. PCF11 encodes a third protein component of yeast cleavage and polyadenylation factor I. *Mol Cell Biol* **17**: 1102-9.

Andrews PD and Stark MJ. 2000. Type 1 protein phosphatase is required for maintenance of cell wall integrity, morphogenesis and cell cycle progression in Saccharomyces cerevisiae. *J Cell Sci* **113**: 507-20.

Aristizibal MJ, Negri GL, Kobor MS. 2015. The RNAPII-CTD maintains genome integrity through inhibition of retrotransposon gene expression and transposition. *PLoS Genet* **11**: e1005608.

Arndt KM and Reines D. 2015. Termination of Transcription of Short Noncoding RNAs by RNA Polymerase II. *Annu Rev Biochem* **84**: 381–404.

Baillat D, Hakimi MA, Naar AM, Shilatifard A, Cooch N, Shiekhattar R. 2005. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**: 265–276.

Bannister AJ and Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell Res* **21**: 381-95.

Barillà D, Lee BA, Proudfoot NJ. 2001. Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisiae. *Proc Natl Acad Sci USA* **98**: 445-50.

Bartolomei MS, Halden NF, Cullen CR, Corden JL. 1988. Genetic analysis of the repetitive carboxyl-terminal domain of the largest subunit of mouse RNA polymerase II. *Mol Cell Biol* **8**: 330-339.

Baskaran R, Chiang GG, Mysliwiec T, Kruch GD, Wang JYJ. 1997. Tyrosine phosphorylation of RNA polymerase II carboxyl-terminal domain by the Abl-related gene product. J Biol Chem 272: 18905–18909.

Baskaran R, Dahmus ME, Wang JY. 1993. Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. *Proc Natl Acad Sci USA* **90**: 11167-71.

Baskaran R, Escobar SR, Wang JY. 1999. Nuclear c-Abl is a COOH-terminal repeated domain (CTD)-tyrosine (CTD)-tyrosine kinase-specific for the mammalian RNA polymerase II: possible role in transcription elongation. *Cell Growth Differ* **10**: 387-96.

Bortvin A and Winston F. 1996. Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* **272**: 1473-6.

Botchkarev VV Jr, Rossio V, Yoshida S. 2014. The budding yeast Polo-like kinase Cdc5 is released from the nucleus during anaphase for timely mitotic exit. *Cell Cycle* **13**: 3260-70.

Brodsky AS and Silver PA. 2000. Pre-mRNA processing factors are required for nuclear export. *RNA* **6**: 1737-49.

Brush GS. 2015. Evidence that histone H1 is dispensable for proper meiotic recombination in budding yeast. *BMC Research Notes* **8**: 275

Bryant JM, Govin J, Zhang L, Donahue G, Pugh BF, Berger SL. 2012. The linker histone plays a dual role during gametogenesis in Saccharomyces cerevisiae. *Mol Cell Biol* **32**: 2771–2783.

Buratowski S. 2003. The CTD code. *Nat Struct Biol* **10**: 679-680.

Burugula BB, Jeronimo C, Pathak R, Jones JW, Robert F, Govind CK. 2014. Histone deacetylases and phosphorylated polymerase II C-terminal domain recruit Spt6 for cotranscriptional histone reassembly. *Mol Cell Biol* **34**: 4115-29.

Bushnell DA, Cramer P, Kornberg RD. 2002. Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 A resolution. *Proc Natl Acad Sci USA* **99**: 1218-1222.

Butler JS and Dent SYR. 2012. Taking stock of transcriptional elongation and histones: increased security through decreased exchange. *Nat Struct Mol Biol* **19**: 863–4.

Calvo O and Manley JL. 2005. The transcriptional coactivator PC4/Sub1 has multiple functions in RNA polymerase II transcription. *EMBO J* **24**: 1009-20.

Carroll KL, Ghirlando R, Ames JM, Corden JL. 2007. Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. *RNA* **13**: 361–373.

Cassart C, Drogat J, Migeot V, Hermand D. 2012. Distinct requirement of RNA polymerase II CTD phosphorylations in budding and fission yeast. *Transcription* **3**: 231–234.

Chapman RD, Heidemann M, Albert TK, Mailhammer R, Flatley A, Meisterernst M, Kremmer E, Eick D. 2007. Transcribing RNA Polymerase II is phosphorylated at CTD residue serine-7. *Science* **318**: 1780-1782.

Chen L, Cai Y, Jin J, Florens L, Swanson SK, Washburn MP, Conaway JW, Conaway RC. 2011. Subunit organization of the human INO80 chromatin remodeling complex: an evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling. *J Biol Chem* **286**: 11283-9.

Chen YZ, Bennett CL, Huynh HM, Blair IP, Puls I, Irobi J, Dierick I, Abel A, Kennerson ML, Rabin BA, Nicholson GA, Auer-Grumbach M, Wagner K, De Jonghe P, Griffin JW, Fischbeck KH, Timmerman V, Cornblath DR, Chance PF. 2004. DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am J Hum Genet* **74**: 1128–1135.

Cho EJ, Kobor MS, Kim M, Greenblatt J, Buratowski S. 2001. Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev* **15**: 3319-29.

Close D, Johnson SJ, Sdano MA, McDonald SM, Robinson H, Formosa T, Hill CP. 2011. Crystal structures of the S. cerevisiae Spt6 core and C-terminal tandem SH2 domain. *J Mol Biol* **408**: 697-713.

Connelly S and Manley JL. 1988. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev* **2**: 440-52.

Conrad NK, Wilson SM, Steinmetz EJ, Patturajan M, Brow DA, Swanson MS, Corden JL. 2000. A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. *Genetics* **154**: 557-71.

Corden JL. 1990. Tails of RNA polymerase II. *Trends in Biochemical Sciences* **15**: 383-7.

DeGennaro CM, Alver BH, Marguerat S, Stepanova E, Davis CP, Bähler J, Park PJ, Winston F. 2013. Spt6 regulates intragenic and antisense transcription, nucleosome positioning, and histone modifications genome-wide in Fission Yeast. *Mol Cell Biol* **33**: 4779–4792.

Dengl S and Cramer P. 2009. Torpedo nuclease Rat1 is insufficient to terminate RNA polymerase II in vitro. *J Biol Chem* **284**; 21270-21279.

Descostes N, Heidemann M, Spinelli L, Schüller R, Maqbool MA, Fenouil R, Koch F, Innocenti C, Gut M, Gut I, Eick D, Andrau JC. 2014. Tyrosine phosphorylation of RNA polymerase II CTD is associated with antisense promoter transcription and active enhancers in mammalian cells. *Elife* **3**: e02105.

Dias JD, Rito T, Triglia ET, Kukalev A, Ferrai C, Chotalia M, Brookes E, Kimura H, Pombo A. 2015. Methylation of RNA polymerase II non-consensus lysine residues marks early transcription in mammalian cells. *eLife* **4**: e11215.

Dichtl B, Blank D, Ohnacker M, Friedlein A, Roeder D, Langen H, Keller W. 2002. A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. *Mol Cell* **10**: 1139-50.

Diebold ML, Loeliger E, Koch M, Winston F, Cavarelli J, Romier C. 2010. Noncanonical Tandem SH2 Enables Interaction of Elongation Factor Spt6 with RNA Polymerase II. *J Biol Chem* **285**: 38389–38398.

Dronamraju R and Strahl BD. 2014. A feed forward circuit comprising Spt6, Ctk1 and PAF regulates Pol II CTD phosphorylation and transcription elongation. *Nucleic Acids Res* **42**: 870-81.

Egloff S. 2012. Role of Ser7 phosphorylation of the CTD during transcription of snRNA genes. *RNA Biol* **9**: 1033–1038.

Egloff S, Dienstbier M, Murphy S. 2012. Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends Genet* **28**: 333-341.

Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S. 2007. Serine-7 of the RNA Polymerase II CTD is specifically required for snRNA gene expression. *Science* **318**: 1777-1779.

Egloff S, Szczepaniak SA, Dienstbier M, Taylor A, Knight S, Murphy S. 2010. The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *J Biol Chem* **285**: 20564-9.

Egloff S [2], Zaborowska J, Laitem C, Kiss T, Murphy S. 2012. Ser7 phosphorylation of the CTD recruits the RPAP2 Ser5 phosphatase to snRNA genes. *Mol Cell* **45**: 111-22.

Eriksson PR, Ganguli D, Nagarajavel V, Clark DJ. 2012. Regulation of histone gene expression in budding yeast. *Genetics* **191**: 7-20.

Filippakopoulos P, Kofler M, Hantschel O, Gish GD, Grebien F, Salah E, Neudecker P, Kay LE, Turk BE, Superti-Furga G, Pawson T, Knapp S. 2008. Structural coupling of SH2-kinase domains links Fes and Abl substrate recognition and kinase activation. *Cell* **134**: 793-803.

Finkel JS, Chinchilla K, Ursic D, Culbertson MR. 2010. Sen1p performs two genetically separable functions in transcription and processing of U5 small nuclear RNA in Saccharomyces cerevisiae. *Genetics* **184**: 107-118.

Francia S. 2015. Non-Coding RNA: Sequence-specific guide for chromatin modification and DNA damage signaling. *Front Genet* **6**: 320.

Gaillard H and Aguilera A. 2014. Cleavage factor I links transcription termination to DNA damage response and genome integrity maintenance in Saccharomyces cerevisiae. *PLoS Genet* **10**: e1004203.

Georgieva M, Roguev A, Balashev K, Zlatanova J, Miloshev. 2012. Hho1p, the linker histone of Saccharomyces cerevisiae, is important for the proper chromatin organization in vivo. *Biochim Biophys Acta* **1819**: 366–374.

Gerber HP, Hagmann M, Seipel K, Georgiev O, West MAL, Litingtung Y, Schaffner W, Corden JL. 1995. RNA polymerase II C-terminal domain required for enhancer-driven transcription. *Nature* **374**: 660-662.

Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. 2003. Global analysis of protein expression in yeast. *Nature* **425**: 737-41.

Ghazy MA, Gordon JMB, Lee SD, Singh BN, Bohm A, Hampsey M, Moore C. 2012. The interaction of Pcf11 and Clp1 is needed for mRNA 3'-end formation and is modulated by amino acids in the ATP-binding site. *Nucleic Acids Res* **40**: 1214-1225.

Greer HL and Shi Y. 2012. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet* **13**: 343-357.

Grzechnik P, Gdula MR, Proudfoot NJ. 2015. Pcf11 orchestrates transcription termination pathways in yeast. *Genes Dev* **29**: 849-61.

Gudipati RK, Villa T, Boulay J, Libri D. 2008. Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nat Struct Mol Biol* **15**: 786-94.

Hanes SD. 2015. Prolyl isomerases in gene transcription. *Biochim Biophys Acta* **1850**:2017-34.

Harlen KM, Trotta KL, Smith EE, Mosaheb MM, Fuchs SM, Churchman LS. 2016. Comprehensive RNA Polymerase II interactomes reveal distinct and varied roles for each phospho-CTD residue. *Cell Rep* **15**: 2147-58.

Henikoff S and Smith MM. 2015. Histone variants and epigenetics. *Cold Spring Harb Perspect Biol* **7**: a019364.

Henry M, Borland CZ, Bossie M, Silver PA. 1996. Potential RNA binding proteins in Saccharomyces cerevisiae identified as suppressors of temperature-sensitive mutations in NPL3. *Genetics* **142**: 103-15.

Heo DH, Yoo I, Kong J, Lidschreiber M, Mayer A, Choi BY, Hahn Y, Cramer P, Buratowski S, Kim M. 2013. The RNA polymerase II C-terminal domain-interacting domain of yeast Nrd1 contributes to the choice of termination pathway and couples to RNA processing by the nuclear exosome. *J Biol Chem* **288**: 36676-90.

Hintermair C, Heidemann M, Koch F, Descostes N, Gut M, Gut I, Fenouil R, Ferrier P, Flatley A, Kremmer E, Chapman RD, Andrau JC, Eick D. 2012. Threonine-4 of mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. *EMBO J* **31**: 2784-97.

Hintermair C, Voss K, Forné I, Heidemann M, Flatley A, Kremmer E, Imhof A, Eick D. 2016. Specific threonine-4 phosphorylation and function of RNA polymerase II CTD during M phase progression. *Sci Rep* **6**: 27401.

Hodko D, Ward T, Chanfreau G. 2016. The Rtr1p CTD phosphatase autoregulates its mRNA through a degradation pathway involving the REX exonucleases. *RNA* **22**: 559-70.

Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, Park JY, Yehia G, Tian B. 2013. Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods* **10**: 133-139.

Houseley J, LaCava J, Tollervey D. 2006. RNA-quality control by the exosome. *Nat Rev Mol Cell Biol* **7**: 529-39.

Howard SC, Budovskaya YV, Chang YW and Herman PK. 2002. The C-terminal Domain of the largest subunit of RNA Polymerase II is required for stationary phase entry and functionally interacts with the Ras/PKA signaling pathway. *JBC* **277**: 19488-19497.

Hsin JP, Li W, Hoque M, Tian B, Manley JL. 2014. RNAP II CTD tyrosine 1 performs diverse functions in vertebrate cells. *eLife* **3**: e02112.

Hsin JP and Manley JL. 2012. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev* **26**: 2119-37.

Hsin JP, Sheth A, Manley JL. 2011. RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3’ end processing. *Science* **334**: 683-686.

Hsin JP, Xiang K, Manley JL. 2014. Function and control of RNA Polymerase II C-Terminal Domain phosphorylation in vertebrate transcription and RNA Processing. *Mol Cell Bio* **34**: 2488–2498.

Hsu PL, Yang F, Smith-Kinnaman W, Yang W, Song JE, Mosley AL, Varani G. 2014. Rtr1 is a dual specificity phosphatase that dephosphorylates Tyr1 and Ser5 on the RNA polymerase II CTD. *J Mol Biol* **426**: 2970-81.

Jamonnak N, Creamer TJ, Darby MM, Schaughency P, Wheelan SJ, Corden JL. 2011. Yeast Nrd1, Nab3, and Sen1 transcriptome-wide binding maps suggest multiple roles in post-transcriptional RNA processing. *RNA* **17**: 2011-2025.

Jasnovidova O, Krejcikova M, Kubicek K, Stefl R. 2017. Structural insight into recognition of phosphorylated threonine-4 of RNA polymerase II C-terminal domain by Rtt103p. EMBO Rep pii: e201643723.

Kelly WG, Dahmus ME, Hart GW. 1993. RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. *J Biol Chem* **268**: 10416-24.

Kim H, Erickson B, Luo W, Seward D, Graber JH, Pollock DD, Megee PC, Bentley DL. 2010. Gene-specific RNA polymerase II phosphorylation and the CTD code. *Nat Struct Mol Biol* **17**: 1279–1286.

Kim M, Krogan NJ, Vasiljeva L, Rando OJ, Nedea E, Greenblatt JF, Buratowski S. 2004. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**: 517-22.

Kim KY and Levin DE. 2011. Mpk1 MAPK association with the Paf1 complex blocks Sen1-mediated premature transcription termination. *Cell* **144**: 745-56.

Kim M, Suh H, Cho EJ, Buratowski S. 2009. Phosphorylation of the yeast Rpb1 C-terminal Domain at serines 2, 5, and 7. *JBC* **284**: 26421-26426.

Komarnitsky P, Cho EJ, Buratowski S. 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev* **14**: 2452-60.

Kong SE, Kobor MS, Krogan NJ, Somesh BP, Søgaard TM, Greenblatt JF, Svejstrup JQ. 2005. Interaction of Fcp1 phosphatase with elongating RNA polymerase II holoenzyme, enzymatic mechanism of action, and genetic interaction with elongator. *J Biol Chem* **280**: 4299-306.

Krishnamurthy S, He X, Reyes-Reyes M, Moore C, Hampsey M. 2004. Ssu72 is an RNA polymerase II CTD phosphatase. *Mol Cell* **14**: 387-394.

Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M, Greenblatt JF, Shilatifard A. 2003. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* **11**: 721-9.

Kubicek K, Cerna H, Holub P, Pasulka J, Hrossova D, Loehr F, Hofr C, Vanacova S, Stefl R. 2012. Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1. *Genes Dev* **26**: 1891-6.

Lemay JF, Marguerat S, Larochelle M, Liu X, van Nues R, Hunyadkürti J, Hoque M, Tian B, Granneman S, Bähler J, Bachand F. 2016. The Nrd1-like protein Seb1 coordinates cotranscriptional 3' end processing and polyadenylation site selection. *Genes Dev* **30**: 1558-72.

Li H, Zhang Z, Wang B, Zhang J, Zhao Y, Jin Y. 2007. Wwp2-Mediated ubiquitination of the RNA Polymerase II large subunit in mouse embryonic pluripotent stem cells. *Mol Cell Bio* **27**: 5296–5305.

Lingner J, Kellermann J, Keller W. 1991. Cloning and expression of the essential gene for poly(A) polymerase from S. cerevisiae. *Nature* **354**: 496-8.

Liu J, Zhang J, Gong Q, Xiong P, Huang H, Wu B, Lu G, Wu J, Shi Y. 2011. Solution structure of tandem SH2 domains from Spt6 protein and their binding to the phosphorylated RNA polymerase II C-terminal domain. *J Biol Chem* **286**: 29218-26.

Liu P, Greenleaf AL, Stiller JW. 2008. The essential sequence elements required for RNAP II carboxyl-terminal domain function in yeast and their evolutionary conservation. *Mol Biol Evol* **25**: 719-27.

Liu P, Kenney JM, Stiller JW, Greenleaf AL. 2010. Genetic organization, length conservation, and evolution of RNA polymerase II carboxyl-terminal domain. *Mol Biol Evol* **27**: 2628-41.

Lu Y, Liu P, James M, Vikis HG, Liu H, Wen W, Franklin A, You M. 2010. Genetic variants cis-regulating Xrn2 expression contribute to the risk of spontaneous lung tumor. Oncogene **29**: 1041-9.

Lunde BM, Reichow SL, Kim M, Suh H, Leeper TC, Yang F, Mutschler H, Buratowski S, Meinhart A, Varani G. 2010. Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **17**: 1195–1201.

Mandal SS, Chu C, Wada T, Handa H, Shatkin AJ, Reinberg D. 2004. Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proc Natl Acad Sci U S A* **101**: 7572-7.

Marina DB, Shankar S, Natarajan P, Finn KJ, Madhani HD. 2013. A conserved ncRNA-binding protein recruits silencing factors to heterochromatin through an RNAi-independent mechanism. *Genes Dev* **27**: 1851–1856.

Maris C, Dominguez C, Allain FH. 2005. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* **272**: 2118-31.

Mayer A, Heidemann M, Lidschreiber M, Schreieck A, Sun M, Hintermair C, Kremmer E, Eick D, Cramer P. 2012. CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. *Science* **336**: 1723-5.

Mayer A, Lidschreiber M, Siebert M, Leike K, Söding J, Cramer P. 2010. Uniform transitions of the general RNA polymerase II transcription complex. *Nature Structural & Molecular Biology* **17**: 1272–1278.

Mbogning J, Pagé V, Burston J, Schwenger E, Fisher RP, Schwer B, Shuman S, Tanny JC. 2015. Functional interaction of Rpb1 and Spt5 C-terminal domains in co-transcriptional histone modification. *Nucleic Acids Res* **43**: 9766–9775.

Mikula M, Skrzypczak M, Goryca K, Paczkowska K, Ledwon JK, Statkiewicz M, Kulecka M, Grzelak M, Dabrowska M, Kuklinska U, Karczmarski J, Rumienczyk I, Jastrzebski K, Miaczynska M, Ginalski K, Bomsztyk K, Ostrowski J. 2016. Genome-wide co-localization of active EGFR and downstream ERK pathway kinases mirrors mitogen-inducible RNA polymerase 2 genomic occupancy. *Nucleic Acids Res* **44**: 10150-10164.

Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, Shilatifard A. 2001. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci USA* **98**: 12902–7.

Mischo HE and Proudfoot NJ. 2013. Disengaging polymerase: Terminating RNA polymerase II transcription in budding yeast. *Biochim Biophys Acta* **1829**: 174-85.

Mishra PK, Ciftci-Yilmaz S, Reynolds D, Au WC, Boeckmann L, Dittman LE, Jowhar Z, Pachpor T, Yeh E, Baker RE, Hoyt MA, D'Amours D, Bloom K, Basrai MA. 2016. Polo kinase Cdc5 associates with centromeres to facilitate the removal of centromeric cohesin during mitosis. *Mol Biol Cell* **27**: 2286-300.

Monneau YR, Soufari H, Nelson CJ, Mackereth CD. 2013. Structure and activity of the peptidyl-prolyl isomerase domain from the histone chaperone Fpr4 towards histone H3 proline isomerization. *J Biol Chem* **288**: 25826-25837.

Morales JC, Richard P, Rommel A, Fattah FJ, Motea EA, Patidar PL, Xiao L, Leskov K, Wu SY, Hittelman WN, Chiang CM, Manley JL, Boothman DA. 2014. Kub5-Hera, the human Rtt103 homolog, plays dual functional roles in transcription termination and DNA repair. *Nucleic Acids Res* **42**: 4996-5006.

Moreira MC, Klur S, Watanabe M, Németh AH, Le Ber I, Moniz JC, Tranchant C, Aubourg P, Tazir M, Schöls L, Pandolfo M, Schulz JB, Pouget J, Calvas P, Shizuka-Ikeda M, Shoji M, Tanaka M, Izatt L, Shaw CE, M'Zahem A, Dunne E, Bomont P, Benhassine T, Bouslam N, Stevanin G, Brice A, Guimarães J, Mendonça P, Barbot C, Coutinho P, Sequeiros J, Dürr A, Warter JM, Koenig M. 2004. Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat Genet* **36**: 225-7.

Morrill SA, Exner AE, Babokhov M, Reinfeld BI, Fuchs SM. 2016. DNA Instability maintains the repeat length of the yeast RNA Polymerase II C-terminal Domain. *J Biol Chem* **291**: 11540-50.

Mosley AL, Pattenden SG, Carey M, Venkatesh S, Gilmore JM, Florens L, Workman JL, Washburn MP. 2009. Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol Cell* **34**: 168-78.

Nedea E, He X, Kim M, Pootoolal J, Zhong G, Canadien V, Hughes T, Buratowski S, Moore CL, Greenblatt J. 2003. Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3'-ends. *J Biol Chem* **278**: 33000-10.

Nelson CJ, Santos-Rosa H, Kouzarides T. 2006. Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell* **126**: 905-16.

Nemec CM, Yang F, Gilmore JM, Hintermair C, Ho YH, Tseng SC, Heidemann M, Zhang Y, Florens L, Gasch AP, Eick D, Washburn MP, Varani G, Ansari AZ. 2017. Different phosphoisoforms of RNA polymerase II engage the Rtt103 termination factor in a structurally analogous manner. Proc Natl Acad Sci USA 114: E3944-E3953.

Ni Z, Olsen JB, Guo X, Zhong G, Ruan ED, Marcon E, Young P, Guo H, Li J, Moffat J, Emili A, Greenblatt JF. 2011. Control of the RNA polymerase II phosphorylation state in promoter regions by CTD interaction domain-containing proteins RPRD1A and RPRD1B. *Transcription* **2**: 237-42.

Nicholson P, Muller B. 2008. Post-transcriptional control of animal histone gene expression: not so different after all. *Mol BioSyst* **4**: 721–725.

Noble CG, Beuth B, Taylor IA. 2007. Structure of a nucleotide-bound Clp1-Pcf11 polyadenylation factor. *Nucleic Acids Res* **35**: 87–99.

Noble CG, Hollingworth D, Martin SR, Ennis-Adeniran V, Smerdon SJ, Kelly G, Taylor IA, Ramos A. 2005. Key features of the interaction between Pcf11 CID and RNA polymerase II CTD. *Nat Struct Mol Biol* **12**: 144-51.

Noble CG, Walker PA, Calder LJ, Taylor IA. 2004. Rna14–Rna15 assembly mediates the RNA-binding capability of Saccharomyces cerevisiae cleavage factor IA. *Nucleic Acids Res* **32**: 3364–3375.

O'Reilly D, Kuznetsova OV, Laitem C, Zaborowska J, Dienstbier M, Murphy S. 2014. Human snRNA genes use polyadenylation factors to promote efficient transcription termination. *Nucleic Acids Res* **42**: 264-75.

Ostapenko D and Solomon MJ. 2003. Budding yeast CTDK-I is required for DNA damage-induced transcription. *Eukaryot Cell* **2**: 274-83.

Paulson AR and Tong L. 2012. Crystal structure of the Rna14-Rna15 complex. *RNA* **18**: 1154-62.

Pearson EL and Moore CL. 2013. Dismantling promoter-driven RNA polymerase II transcription complexes in vitro by the termination factor Rat1. *J Biol Chem* **288**: 19750-9.

Pearson G, Robinson F, Gibson TB, Xu B, Karandikar M, Berman K, Cobb MH. 2011. Mitogen-Activated Protein (MAP) kinase pathways: Regulation and physiological functions. *Endocrine Reviews* **22**: 153–83.

Pirngruber J, Shchebet A, Schreiber L, Shema E, Minsky N, Chapman RD, Eick D, Aylon Y, Oren M, Johnsen SA. 2009. CDK9 directs H2B monoubiquitination and controls replication-dependent histone mRNA 3'-end processing. *EMBO Rep* **10**: 894-900.

Porrua O, Hobor F, Boulay J, Kubicek K, D'Aubenton-Carafa Y, Gudipati RK, Stefl R, Libri D. 2012. In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. *EMBO J* **31**: 3935-48.

Porrua O and Libri D. 2013. A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. *Nat Struct Mol Biol* **20**: 884–891.

Porrua O and Libri D. 2015. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat Rev Mol Cell Biol* **16**: 190–202.

Qiu H, Hu C, Hinnebusch AG. 2009. Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. *Mol Cell* **33**: 752-62.

Quenet D and Dalal Y. The CENP-A nucleosome: a dynamic structure and role at the centromere. *Chromosome Res.* **20**: 465-479.

Ramakrishnan S, Pokhrel S, Palani S, Pflueger C, Parnell TJ, Cairns BR, Bhaskara S, Chandrasekharan MB. 2016. Counteracting H3K4 methylation modulators Set1 and Jhd2 co-regulate chromatin dynamics and gene transcription. *Nat Comm* **7**: 11949.

Ranuncolo SM, Ghosh S, Hanover JA, Hart GW, Lewis BA. 2012. Evidence of the involvement of O-GlcNAc-modified human RNA Polymerase II CTD in transcription in vitro and in vivo. *J Biol Chem* **287**: 23549–61.

Richard P, Feng S, Manley JL. 2013. A SUMO-dependent interaction between Senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage. *Genes Dev* **27**: 2227-32.

Richard P and Kiss T. 2006. Integrating snoRNP assembly with mRNA biogenesis. *EMBO Rep* **7**: 590-2.

Rosonina E, Yurko N, Li W, Hoque M, Tian B, Manley JL. 2014. Threonine-4 of the budding yeast RNAP II CTD couples transcription with Htz1-mediated chromatin remodeling. *Proc Natl Acad Sci USA* **111**: 11924-31.

Rossetto D, Avvakumov N, Côté J. 2012. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. Epigenetics **7**: 1098-108.

Saberianfar R, Cunningham-Dunlop S, Karagiannis J. 2011. Global gene expression analysis of fission yeast mutants impaired in Ser-2 phosphorylation of the RNA pol II carboxy terminal domain. *PloS One* **6**: e24694.

Sadowski M, Dichtl B, Hübner W, Keller W. 2003. Independent functions of yeast Pcf11p in pre-mRNA 3' end processing and in transcription termination. *EMBO J* **22**: 2167-2177.

Sainsbury S, Bernecky C, Cramer P. 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**: 129–143.

Schlackow M, Nojima T, Gomes T, Dhir A, Carmo-Fonseca M, Proudfoot NJ. 2017. Distinctive patterns of transcription and RNA processing for human lincRNAs. Mol Cell 65: 25-38.

Schreieck A, Easter AD, Etzold S, Wiederhold K, Lidschreiber M, Cramer P, Passmore LA. 2014. RNA polymerase II termination involves C-terminal-domain tyrosine dephosphorylation by CPF subunit Glc7. *Nat Struct Mol Biol* **21**: 175-9.

Schroeder SC, Schwer B, Shuman S, Bentley D. 2000. Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes Dev* **14**: 2435-40.

Schüller R, Forné I, Straub T, Schreieck A, Texier Y, Shah N, Decker TM, Cramer P, Imhof A, Eick D. 2016. Heptad-specific phosphorylation of RNA polymerase II CTD. *Mol Cell* **61**: 305–314.

Schwer B, Bitton DA, Sanchez AM, Bähler J, Shuman S. 2014. Individual letters of the RNA polymerase II CTD code govern distinct gene expression programs in fission yeast. *PNAS* **111**: 4185–4190.

Schwer B, Sanchez AM, Shuman S. 2012. Punctuation and syntax of the RNA polymerase II CTD code in fission yeast. *Proc Natl Acad Sci* **109**: 18024–18029.

Schwer B, Sanchez AM, Shuman S. 2015. RNA polymerase II CTD phospho-sites Ser5 and Ser7 govern phosphate homeostasis in fission yeast. RNA 21: 1770–1780.

Schwer B, Shuman S. 2011. Deciphering the RNA polymerase II CTD code in fission yeast. *Mol Cell* **43**: 311–318.

Shandilya J, Roberts SG. 2012. The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. *Biochem Biophys Acta* **1819**:391-400.

Sikorski TW, Buratowski S. 2009. The basal initiation machinery: beyond the general transcription factors. *Curr Opin Cell Biol* **21**: 344-351.

Sims RJ 3rd, Rojas LA, Beck D, Bonasio R, Schüller R, Drury WJ 3rd, Eick D, Reinberg D. 2011. The C-terminal domain of RNA polymerase II is modified by site-specific methylation. *Science* **332**: 99-103.

Skaar JR, Ferris AL, Wu X, Saraf A, Khanna KK, Florens L, Washburn MP, Hughes SH, Pagano M. 2015. The Integrator complex controls the termination of transcription at diverse classes of gene targets. *Cell Res* **25**: 288–305.

Skourti-Stathaki K, Proudfoot NJ, Gromak N. 2011. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell* **42**: 794–805.

South PF, Harmeyer KM, Serratore ND, Briggs SD. 2013. H3K4 methyltransferase Set1 is involved in maintenance of ergosterol homeostasis and resistance to Brefeldin A. *Proc Natl Acad Sci U S A* **110**: E1016-25.

Spain MM, Ansari SA, Pathak R, Palumbo MJ, Morse RH, Govind CK. 2014. The RSC complex localizes to coding sequences to regulate Pol II and histone occupancy. *Mol Cell* **56**: 653-66.

Srividya I, Tirupataiah S, Mishra K. 2012. Yeast transcription termination factor Rtt103 functions in DNA damage response. *PLoS One* **7**: e31288.

St Amour CV, Sansó M, Bösken CA, Lee KM, Larochelle S, Zhang C, Shokat KM, Geyer M, Fisher RP. 2012. Separate domains of fission yeast Cdk9 (P-TEFb) are required for capping enzyme recruitment and primed (Ser7-phosphorylated) Rpb1 carboxyl-terminal domain substrate recognition. *Mol Cell Biol* **32**: 2372-83.

Steinmetz EJ, Conrad NK, Brow DA, Corden JL. 2001. RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature* **413**: 327-331.

Sterner DE, Lee JM, Hardin SE, Greenleaf AL. 1995. The yeast carboxyl-terminal repeat domain kinase CTDK-I is a divergent cyclin-cyclin-dependent kinase complex. *Mol Cell Biol* **15**: 5716-24.

Stiller JW, Mcconaughy BL, Hall BD. 2000. Evolutionary complementation for polymerase II CTD function. *Yeast* **16**: 57–64.

Strich R and Cooper KF. 2014. The dual role of cyclin C connects stress regulated gene expression to mitochondrial dynamics. *Microb Cell* **1**: 318–324.

Suh H, Ficarro SB, Kang U, Chun Y, Marto JA, Buratowski S. 2016. Direct analysis of phosphorylation sites on the Rpb1 C-Terminal Domain of RNA Polymerase II. *Mol Cell* **61**: 297–304.

Sun M, Larivière L, Dengl S, Mayer A, Cramer P. 2010. A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA Polymerase II C-terminal repeat Domain (CTD). *J Biol Chem* **285**: 41597-41603.

Takagaki Y and Manley JL. 1998. Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation. *Mol Cell* **2**: 761-71.

Tanny JC. 2014. Chromatin modification by the RNA Polymerase II elongation complex. *Transcription* **5**: e988093.

Terzi N, Churchman LS, Vasiljeva L, Weissman J, Buratowski S. 2011. H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway. *Mol Cell Biol* **31**: 3569–3583.

Tietjen JR, Zhang DW, Rodríguez-Molina JB, White BE, Akhtar MS, Heidemann M, Li X, Chapman RD, Shokat K, Keles S, Eick D, Ansari AZ. 2010. Chemical-genomic dissection of the CTD code. *Nat Struct Mol Biol* **17**: 1154-61.

Tsao DC, Park NJ, Nag A, Martinson HG. 2012. Prolonged α-amanitin treatment of cells for studying mutated polymerases causes degradation of DSIF160 and other proteins. *RNA* **18**: 222-9.

Tu J and Carlson M. 1994. The GLC7 type 1 protein phosphatase is required for glucose repression in Saccharomyces cerevisiae. *Mol Cell Biol* **14**: 6789-96.

Tudek A, Porrua O, Kabzinski T, Lidschreiber M, Kubicek K, Fortova A, Lacroute F, Vanacova S, Cramer P, Stefl R, Libri D. 2014. Molecular basis for coordinating transcription termination with noncoding RNA degradation. *Mol Cell* **55**: 467-81.

Vasiljeva L and Buratowski S. 2006. Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. *Mol Cell* **21**: 239-248.

Vasiljeva L, Kim M, Mutschler H, Buratowski S, Meinhart A. 2008. The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **15**: 795-804.

Verdin E and Ott M. 2015. 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat Rev Mol Cell Biol* **16**: 258-264.

Wagner EJ and Carpenter PB. 2012. Understanding the language of Lys36 methylation at histone H3. *Nat Rev Mol Cell Biol* **13**: 115-26.

Wang J, Takagaki Y, Manley JL. 1996. Targeted disruption of an essential vertebrate gene: ASF/SF2 is required for cell viability. *Genes Dev* **10**: 2588-99.

Weiner A, Chen HV, Liu CL, Rahat A, Klien A, Soares L, Gudipati M, Pfeffner J, Regev A, Buratowski S, Pleiss JA, Friedman N, Rando OJ. 2012. Systematic dissection of roles for chromatin regulators in a yeast stress response. *PLoS Biol* **10**: e1001369.

West ML and Corden JL. 1995. Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. *Genetics* **140**: 1223-33.

West S, Gromak N, Proudfoot NJ. 2004. Human 5'-3' exonuclease XRN2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* **432**: 522-525.

West S and Proudfoot NJ. 2008. Human Pcf11 enhances degradation of RNA polymerase II-associated nascent RNA and transcriptional termination. *Nucleic Acids Res* **36**: 905-14.

Winsor TS, Bartkowiak B, Bennett CB, Greenleaf AL. 2013. A DNA damage response system associated with the phosphoCTD of elongating RNA polymerase II. *PLoS One* **8**: e60909.

Xiang K, Manley JL, Tong L. 2012. An unexpected binding mode for a Pol II CTD peptide phosphorylated at Ser7 in the active site of the CTD phosphatase Ssu72. *Genes Dev* **26**: 2265–2270.

Xiang K, Manley JL, Tong L. 2012 [2]. The yeast regulator of transcription protein Rtr1 lacks an active site and phosphatase activity. *Nat Comm* **3**: 946.

Xu YX, Hirose Y, Zhou XZ, Lu KP, Manley JL. 2003. Pin1 modulates the structure and function of human RNA polymerase II. *Genes Dev* **17**: 2765-76.

Yang F, Hsu P, Lee SD, Yang W, Hoskinson D, Xu W, Moore C, Varani G. 2017. The C terminus of Pcf11 forms a novel zinc-finger structure that plays an essential role in mRNA 3'-end processing. *RNA* **23**: 98-107.

Youdell ML, Kizer KO, Kisseleva-Romanova E, Fuchs SM, Duro E, Strahl BD, Mellor J. 2008. Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. *Mol Cell Biol* **28**: 4915-26.

Yuryev A, Patturajan M, Litingtung Y, Joshi RV, Gentile C, Gebara M, Corden JL. 1996. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc Natl Acad Sci USA* **93**: 6975-6980.

Zeidan Q and Hart GW. 2010. The intersections between O-GlcNAcylation and phosphorylation: implications for multiple signaling pathways. *Journal of Cell Science* **123**: 13–22.

Zhang DW, Mosley AL, Ramisetty SR, Rodríguez-Molina JB, Washburn MP, Ansari AZ. 2012. Ssu72 phosphatase-dependent erasure of phospho-Ser7 marks on the RNA Polymerase II C-terminal Domain is essential for viability and transcription termination. *J Biol Chem* **287**: 8541–8551.

Zhang Y, Zhang M, Zhang Y. 2011. Crystal structure of Ssu72, an essential eukaryotic phosphatase specific for the C-terminal domain of RNA polymerase II, in complex with a transition state analogue. *Biochem J* **434**: 435-44.

Zhang Z, Fu J, Gilmour DS. 2005. CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. *Genes Dev* **19**: 1572–1580.

Zhou P, Wu E, Alam HB, Li Y. 2014. Histone cleavage as a mechanism for epigenetic regulation: Current insights and perspectives. *Curr Mol Med* **14**: 1164–1172.

Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, Smith D, Gerstein M, Reed MA, Snyder M. 2000. Analysis of yeast protein kinases using protein chips. Nat Genet 26: 283-9.

Zimmerman WC and Erikson RL. 2007. Finding Plk3. *Cell Cycle* **6**: 1314-8.

Figure 1 – Ser7

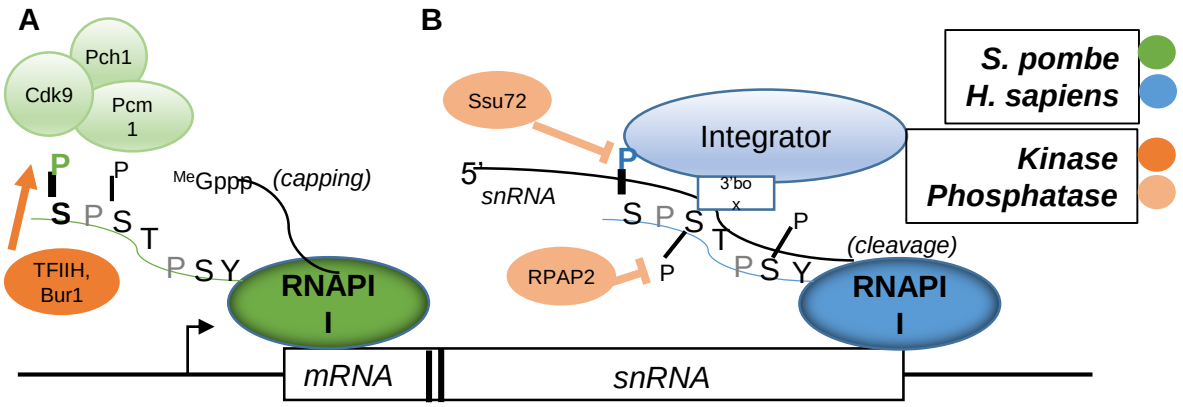


Figure 1: Ser7 facilitates mRNA capping in *S. pombe* and snRNA processing in *H. sapiens*. (A) In *S. pombe*, Cdk9 (with cyclin partner Pcm1) binds to CTD Ser7P, and enables the recruitment of 5' cap methyltransferase Pcm1 to a Ser5P CTD. Ser7 is phosphorylated by TFIIH subunit Cdk7/Mcs6, as well as by Bur1. (B) In *H. sapiens*, the 12-subunit Integrator complex preferentially recognizes a Ser7P-Ser2P CTD at the 3' end of snRNA genes. Ser7P recruits RPAP2 to the CTD, which dephosphorylates Ser5P, and Integrator is able to recognize both the Ser7P-Ser2P CTD as well as the 3'box of the snRNA, enabling cleavage and processing of the snRNA. TFIIH subunit Cdk7 phosphorylates Ser7, while phosphatase Ssu72 has been shown to dephosphorylate Ser7P.

Figure 2 – Thr4

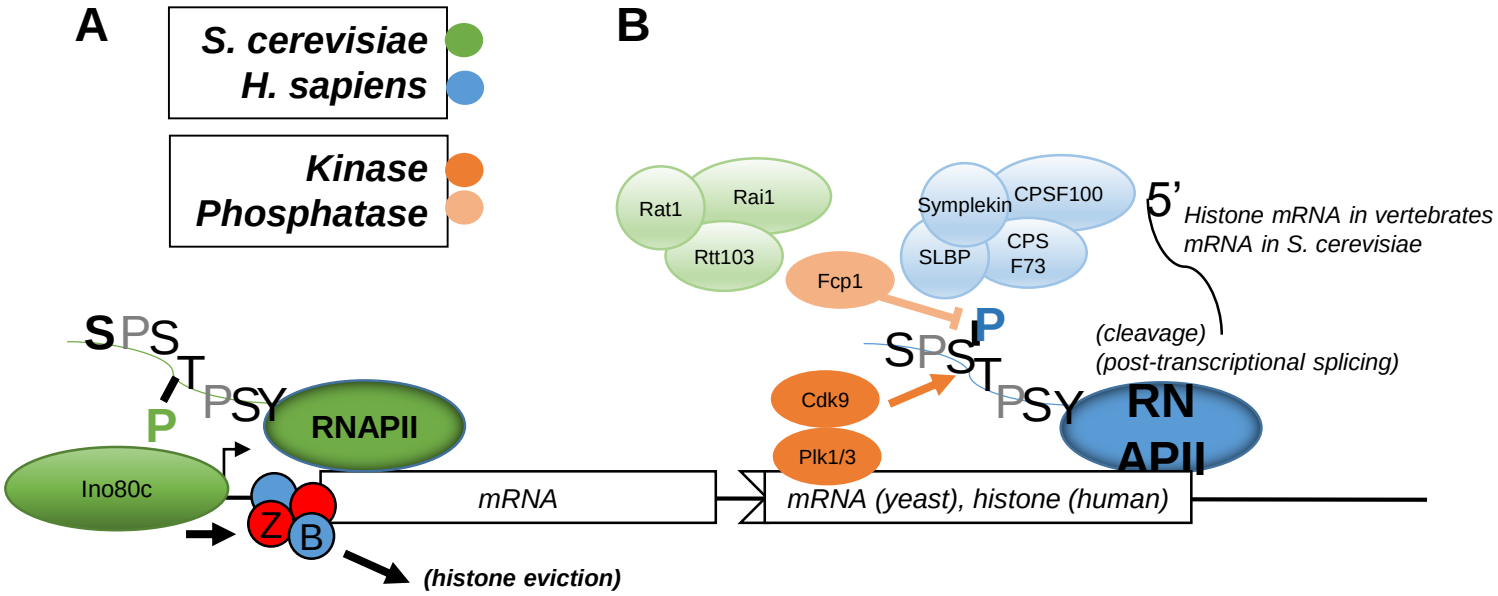


Figure 2: Divergent functions of Thr4 between *S. cerevisiae* and *H. sapiens*. (A) In *S. cerevisiae*, Thr4P has been shown to recruit the multi-subunit Ino80 chromatin remodeling complex to certain classes of promoters (see text), which evicts the promoter-proximal H2A.Z (red)/H2B (blue) histone dimers. Upon eviction of these dimers and replacement with H2A/H2B dimers, transcription of these genes is enabled. (B) In *H. sapiens*, histone mRNA synthesis requires Thr4 for efficient 3' processing. Thr4 is required for recruitment of stem-loop binding protein (SLBP) and CPSF100, required along with other canonical 3' processing factors such as Symplekin and CPSF73, to replication-dependent histone genes. In *S. cerevisiae*, histone mRNA 3' ends are formed identically to all other mRNAs, and evidence shows that instead Thr4P plays a role in post-transcriptional splicing and Rtt103 (and Rat1/Rai1) recruitment during termination. In *H. sapiens*, Cdk9 and Plk1/Plk3 have been shown to phosphorylate Thr4, the later perhaps specifically in M phase, and Thr4P is dephosphorylated by Fcp1. Neither the kinase(s) nor phosphatase that act on Thr4 are known in *S. cerevisiae*.

Figure 3 – Tyr1

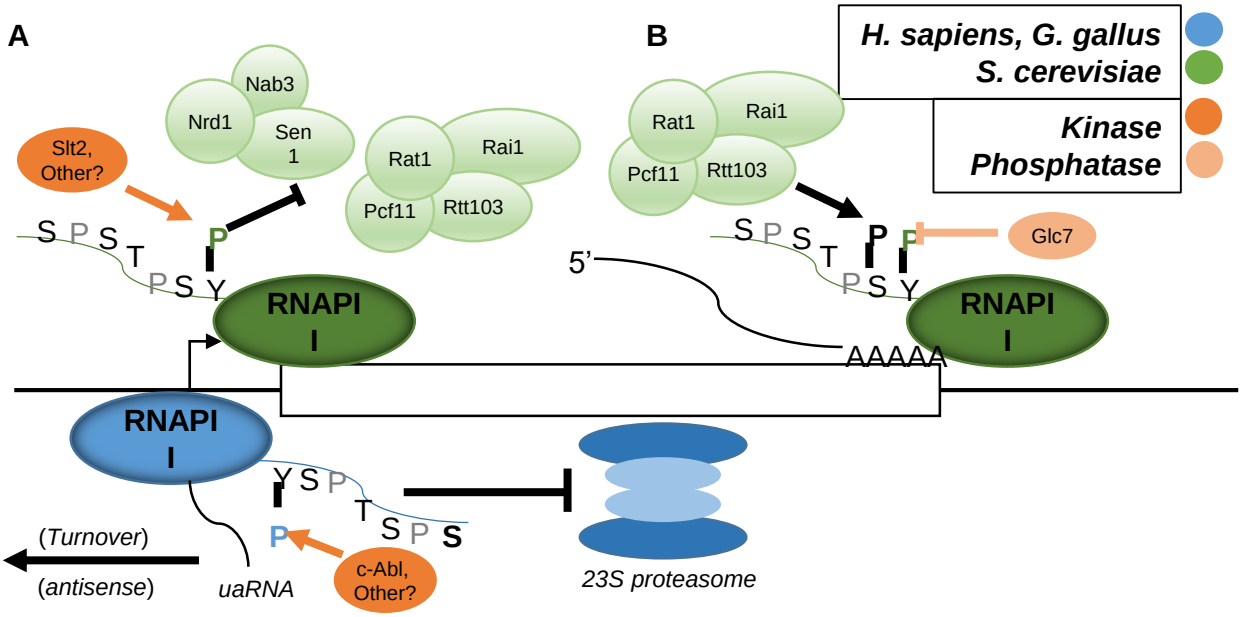


Figure 3: Tyr1 has multiple functions across species. (A) At the 5' ends of genes, Tyr1 in vertebrates is phosphorylated, likely by c-Abl, and Tyr1P is important for efficient turnover of 5' upstream antisense RNAs (uaRNAs). Tyr1 phosphorylation also enhances stability and prevents degradation of the unstructured CTD by the 23S proteasome in vertebrate cells. In *S. cerevisiae*, this function has not been described, and Tyr1 phosphorylation (by Slt2 or other unidentified kinases) is associated with antitermination, as it prevents efficient binding of Nrd1 (in yeast) or Rtt103 (yeast and humans) to a Ser5P or Ser2P CTD, respectively. (B) At 3' ends of genes, CTD Tyr1P is dephosphorylated by Glc7, a subunit of cleavage and polyadenylation factor (CPF). This leads to Ser2P-CTD recognition by Pcf11 and Rtt103/Rat1/Rai1, which facilitates efficient termination. This function appears conserved from yeast to human.

Chapter 2

Threonine-4 of the budding yeast RNAP II CTD couples transcription with Htz1-mediated chromatin remodeling

Emanuel Rosonina[1], Nathan Yurko[1], Wencheng Li[2], Mainul Hoque[2], Bin Tian[2], and James L. Manley[1]

[1]Department of Biological Sciences, Columbia University, New York, NY 10027, USA

[2]Department of Microbiology, Biochemistry and Molecular Genetics, Rutgers New Jersey Medical School, Newark, New Jersey, USA

Abstract

The C-terminal domain (CTD) of the largest subunit of RNA polymerase II (RNAP II) consists of repeated YSPTSPS heptapeptides and connects transcription with cotranscriptional events. Threonine-4 (Thr4) of the CTD repeats has been shown to function in histone mRNA 3'-end processing in chicken cells and in transcriptional elongation in human cells. Here, we demonstrate that, in budding yeast, Thr4, although dispensable for growth in rich media, is essential in phosphate-depleted or galactose-containing media. Thr4 is required to maintain repression of phosphate-regulated (PHO) genes under normal growth conditions and for full induction of *PHO5* and the galactose-induced *GAL1* and *GAL7* genes. We identify genetic links between Thr4 and the histone variant Htz1 and show that Thr4, as well as the Ino80 chromatin remodeler, is required for activation-associated eviction of Htz1 specifically from promoters of the Thr4-dependent genes. Our study uncovers a connection between transcription and chromatin remodeling linked by Thr4 of the CTD.

Introduction

In eukaryotic cells, RNA polymerase II (RNAP II) transcribes all protein coding genes as well as a number of noncoding RNA genes, including those for most snRNAs and microRNAs. Expression of RNAP II-transcribed genes is highly regulated, ensuring that the timing of production, quantity, and structure of the RNAs generated is appropriate. Much of this regulation occurs on the genes themselves, whose accessibility is limited by nucleosomes that can form densely packed chromatin (reviewed in refs. 1 and 2). A significant amount of regulation, however, targets RNAP II itself. Specifically, the C-terminal domain (CTD) of Rpb1, the largest subunit of RNAP II,

is subject to a multitude of phosphorylation and dephosphorylation events that mark different steps of the RNAP II transcription cycle (recently reviewed in refs. 3–5). Although many of these modifications occur generally, some appear to play critical roles in expression of specific genes (6–8). The CTD consists of a tandemly repeated hepta-peptide, whose consensus sequence, Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7, is conserved in yeast, which harbors 26 repeats, and vertebrates, with 52 repeats. The emerging picture places this unusual structure in a highly important role in coordinating transcription by RNAP II with other events in the generation and maturation of RNAs.

Since its discovery, numerous studies have focused on the role of the entire CTD and the function of individual amino acids within its repeats. At promoters, the unphosphorylated CTD interacts directly with subunits of the Mediator complex (9), which is required for Mediator's role in stimulating transcription (10). Additionally, early studies determined that the CTD is required for full activation by some promoter-bound activators (11–13) whereas it was later shown that the CTD can mediate the influence of activators on pre-mRNA processing (14–16).

Several genome-wide analyses uncovered a pattern of CTD phosphorylation that occurs once RNAP II clears promoters (e.g., refs. 17 and 18). Phosphorylation of Ser5 occurs near the 5' end of genes where it enables recruitment of the capping machinery to the nascent pre-mRNA (19, 20), as well as recruitment of Set1, a COMPASS-associated methyltransferase that modifies histone H3-Lys4 around the 5' end of transcriptionally active genes (21, 22). Toward the 3' end of genes, Ser5 phosphorylation decreases, Ser2 phosphorylation increases, and CTD repeats that are dually phosphorylated on Ser5 and Ser2 can bind the Set2 methyltransferase, thereby

targeting histone H3 for trimethylation on Lys36 (23). Furthermore, Ser2 phosphorylation functions in the recruitment of splicing, 3' processing, and termination factors (reviewed in refs. 4 and 5). Like Ser5, Ser7 is phosphorylated near the beginning of the transcription cycle, and, although it is detected on both protein-coding and noncoding genes, mutation of Ser7 to Ala in human Rpb1 caused defects in expression specifically of snRNA genes (7). These studies demonstrate that the CTD not only participates in the production and processing of RNA but also leaves behind chromatin modifications that demarcate regions of transcriptionally active genes.

We previously characterized the role of threonine-4 (Thr4) of the CTD repeats by generating chicken DT40 cells that express a mutant form of Rpb1 in which each Thr4 was replaced with Val (8). Cells expressing the Thr-Val (T4V) form of the CTD were not able to survive past 24 h, likely due to defects that were detected in 3'-end processing specifically of histone pre-mRNAs. In metazoans, histone mRNA 3' processing involves a mechanism that is distinct from the typical polyadenylation of other pre-mRNAs, whose processing and expression were not affected by the Thr4 mutation. In another study, ChIP-seq analysis determined that Thr4 phosphorylation increases toward the 3' end of RNAP II-transcribed genes and that Thr-to-Ala mutation caused defects in transcription elongation in a human cell line (24). In budding yeast, histone pre-mRNAs are not processed by a distinct mechanism, leading to the question of whether Thr4 plays gene-specific and/or general roles in transcription in that organism. Thr4 mutation in fission yeast was shown to be without effect on growth (25, 26).

Here, we address the role of Thr4 in *Saccharomyces cerevisiae* by characterizing a strain that expresses a T4V mutant form of Rpb1. The mutant strain is viable and

shows no general transcriptional defect in normal growth conditions but is unable to grow in low phosphate or galactose-containing medium. Linking Thr4 with phosphate and galactose metabolism, we found that T4V cells are unable to fully repress phosphate-regulated (PHO) genes under normal growth conditions and to fully activate *PHO5* and the galactose-inducible *GAL1* and *GAL7* genes under their respective inducing conditions. We identify genetic links between T4V and genes encoding the SWR1 and INO80 chromatin remodeling complex subunits, as well as the gene encoding the target of these complexes, the histone variant Htz1. Connecting our observations, we demonstrate that eviction of Htz1 from promoter-proximal nucleosomes, a process associated with gene activation, is defective in T4V cells, specifically for the genes requiring Thr4 for activation, and we implicate the Ino80 complex in Thr4-mediated loss of Htz1 from these genes. Our study, therefore, links Thr4 of the yeast CTD with promoter-associated chromatin remodeling necessary for control of activated transcription.

Results

Thr4 Is Phosphorylated on RNAP II Transcribed Genes but Is Not Essential for Viability in Rich Medium

In our previous study, we determined that phosphorylated Thr4 (P-Thr4) is present on Rpb1 in chicken and human cells, and in budding yeast (8). To characterize Thr4 phosphorylation in yeast, we first conducted chromatin fractionation using yeast extracts and found by Western blot (WB) with a P-Thr4 specific antibody that Thr4 is phosphorylated specifically on chromatin-associated RNAP II (Fig. 1A), suggesting that Thr4 functions during transcription. To explore this possibility, we constructed a yeast

strain in which Thr4 on each repeat of the CTD was replaced by Val, analogous to what was done in DT40 cells (8). This strain, named T4V, expresses only Rpb1 containing 24 repeats of the sequence YSPVSPS. In addition to replacing Thr with Val, this mutant also eliminated the naturally occurring variation found in some repeats of the CTD (27). To account for any effects that might result from this loss of nonconsensus residues instead of the Thr-Val switch, we constructed a control strain that expressed an Rpb1 with 26 consensus CTD repeats, named CONS. Fig. 1B confirms by WB that Thr4 was phosphorylated in the WT and CONS strains, but not in the T4V strain, and that Rpb1 levels in T4V, CONS, and WT strains were equivalent (SI Appendix, Fig. S5A). Comparison of growth of the three strains showed that T4V displayed a transient and modest growth defect, compared with WT and CONS, on both rich and synthetic complete (SC) medium and that this defect was exacerbated at low temperatures (18 °C) (Fig. 1C). Chromatin immunoprecipitation (ChIP) analysis showed that P-Thr4 was detected on constitutively active and activated inducible genes and that the T4V mutation did not significantly affect RNAP II levels on constitutive genes (SI Appendix, Fig. S1 A–D).

Thr4 Is Required for Repression of Phosphate-Regulated Genes in Normal Medium

Our observation that Thr4 is not essential for growth indicates that it is not generally required for RNAP II transcription. However, to determine whether Thr4 is involved in transcription of specific genes, we first performed global mRNA sequencing (RNA-seq) in WT and T4V strains grown in SC medium. A summary of the most significantly up- and down-regulated genes is shown in Fig. 2A, and the full results are

in Dataset S1. Seventeen genes displayed an average increased expression of at least threefold (log2 > ~1.6) whereas nine genes showed an average drop in expression of at least threefold (log2 < ~1.6). A diagram of WT and T4V RNA-seq read distributions over the chromosomal loci for up-regulated genes PHO84 (with unaffected flanking genes) and PHO5, and down-regulated gene PTC2, is shown in Fig. 2B. Interestingly, approximately half of the up-regulated genes are involved in phosphate metabolism (Fig. 2A, indicated by asterisks). Specifically, these genes, hereafter referred to as PHO genes, are also up-regulated under growth conditions of low inorganic phosphate (Pi) levels (28). RT-PCR analysis on a subset of the regulated genes validated the RNA-seq results and confirmed that the T4V mutation resulted in up-regulation of PHO genes (SI Appendix, Fig. S2).

To determine whether these changes in gene expression were specific to the T4V mutation and not due to the loss of nonconsensus CTD repeats, we performed an additional RNA-seq analysis comparing mRNAs from CONS and WT strains. This analysis showed no similarity with the T4V vs. WT RNA-seq analysis (SI Appendix, Fig. S3). Although PHO84 mRNA levels were approximately threefold higher in CONS than in WT (log2 of 1.68), this increase was dramatically less than in T4V (>1,000-fold, log2 of 10.8), and no derepression of other PHO genes was observed in the CONS strain (Fig. 2A and Dataset S2). Some PHO-regulated genes, however, also appeared up-regulated in the CONS strain in the RT-PCR validation (SI Appendix, Fig. S2). An explanation for this apparent discrepancy stems from the fact that the RT-PCR analysis was performed with random-primed RNA, which detected polyA− as well polyA+ transcripts whereas the RNA-seq analysis detected only polyA+ RNAs. Thus, a

parsimonious interpretation of these results is that loss of nonconsensus CTD residues resulted in production of polyA− transcripts, which could be sense and/or antisense, from some PHO genes, whereas mutation specifically of Thr4 led to up-regulation of polyA+ sense transcripts from all PHO genes analyzed. Furthermore, RNA-seq of a strain expressing Rpb1 with a mutant CTD composed entirely of otherwise consensus repeats in which all Tyr1 residues were changed to Phe (Y1F) revealed no up-regulation of the PHO genes. Our results, therefore, demonstrate that the CTD is involved in repressing PHO gene transcription and that Thr4 functions specifically in preventing the production of polyadenylated PHO mRNAs during normal growth.

To extend the above results, we examined growth of the T4V strain in low Pi medium. Strikingly, whereas WT and CONS strains grew normally on this medium, T4V was inviable (Fig. 2C). T4V sensitivity to low Pi medium was observed in both S288C and W303 background strains. Although both the RNA-seq and RT-PCR analyses showed significant derepression of PHO-regulated genes in the T4V strain, derepressed transcription levels for *PHO5* and *PHO84* were still significantly lower than observed during normal induction of these genes in low Pi conditions (Fig. 2D). Our observations, therefore, indicate that, although Thr4 is required for normal expression of only a small number of genes under normal growth conditions, it is necessary for repression of PHO-regulated genes and plays an essential role in the cellular response to low phosphate levels.

Thr4 Is Genetically Linked to HTZ1 and the SWR1 and INO80 Chromatin Remodeling Complexes

To identify pathways that mediate the effects of Thr4 on growth and transcription, we carried out genome-scale yeast genetic interaction assays [synthetic genetic array (SGA)] (29) with the T4V and CONS strains. Genes whose deletion in combination with the T4V mutation (but not with the CONS mutation) resulted in the most dramatic growth defect in both of two independent SGA analyses are listed in SI Appendix, Fig. S4, and the subset of these genes that is involved in transcription and chromatin regulation is indicated in Fig. 3A. Notably, several components of the SWR1 chromatin remodeling complex were among the strongest genetic interactions with the T4V mutation, which we confirmed by comparing the growth of independently derived double-mutant strains with single-mutant strains (Fig. 3B). The SWR1 complex replaces conventional histone H2A with the variant Htz1 (also known as H2A.Z) in nucleosomes (30–32), specifically those associated with promoters of transcriptionally inactive genes (33–35). Additionally, the gene encoding Ies2, a subunit of another chromatin remodeling complex, the INO80 complex (INO80c), showed a strong genetic interaction with T4V (Fig. 3A) (most other genes encoding INO80c subunits are essential and were therefore not represented in the screen). INO80c can exchange Htz1/H2B dimers for H2A/H2B in in vitro assays and is required for normal localization of Htz1 across the genome, suggesting that it functions in cells to remove Htz1 from activated promoters (36).

The *HTZ1* gene was also a strong hit in the SGA screen. Consistent with this genetic interaction, *HTZ1* and the T4V mutation showed a synthetic lethal interaction in multiple independent tetrad dissection analyses (Fig. 3C). T4V spores derived from *HTZ1/htz1Δ WT-RPB1/T4V* double heterozygotes displayed a more severe growth

defect than normal after dissection (compare Fig. 3C, Upper with Fig. 1C; and see also Fig. 5B) although, like other T4V strains, the defect was transient and T4V approached WT growth after ~48 h (Fig. 3C, Lower). Additionally, attempts to delete *HTZ1* in a haploid T4V strain by homologous recombination produced no viable transformants. In another approach, we deleted *HTZ1* in WT, CONS, or T4V strains harboring an *HTZ1* expression plasmid (with a URA3 marker) and then eliminated the plasmid by growth on 5-FOA. Only the T4V *htz1Δ* background showed no growth on 5-FOA (Fig. 3D), supporting the synthetic genetic interaction between *HTZ1* and T4V. These results together strongly implicate Thr4 in Htz1-related pathways.

Thr4 Is Required for Expression of PHO5, GAL1, and GAL7 and for Eviction of Htz1 from Their Promoters During Induction

Htz1 is typically found within the two nucleosomes that flank the nucleosome-free region at RNAP II promoters (34, 37), and, although it is not required for maintaining a repressed state, its removal from these nucleosomes stimulates gene activation (33, 35). Correspondingly, genomic analyses localized Htz1 primarily at repressed gene promoters, and its occupancy is inversely proportional to transcription levels (33, 35, 37, 38). Relating Htz1 function with our observed role for Thr4 in PHO gene regulation, a previous study (39) demonstrated that *HTZ1* acts synthetically with genes encoding the SWI/SNF chromatin-remodeling complex to activate expression of *PHO5*, which encodes the cell's major acid phosphatase (40). We compared growth of *htz1*Δ and T4V cells on low Pi medium and found that the two strains displayed strong growth defects, indicating that both Thr4 and Htz1 indeed play critical roles in the PHO pathway (Fig. 4A). We then determined whether Thr4 and Htz1 are needed for expression of genes

induced during low Pi conditions and found that both T4V and *htz1*Δ strains showed reduced expression of *PHO5* whereas induction of another PHO gene, *PHO84*, was not affected (Fig. 4B).

We also tested for growth defects on other media in which gene induction is necessary for growth. We found that T4V was also highly sensitive to medium containing galactose as the sole carbon source but grew normally on amino acid-depleted or high osmolarity media (Fig. 4A). Low Pi and galactose sensitivity was specific to the T4V mutation because the CONS and strain did not show specific growth defects on these media (Fig. 4A). Galactose sensitivity correlated with reduced expression of the galactose-induced *GAL1* and *GAL7* genes in T4V (Fig. 4B) whereas expression of the amino acid starvation-induced *ARG1* and high osmolarity-induced *STL1* genes was not affected. Notably, a previously described strain expressing Rpb1 with a CTD truncated to ~13 repeats showed normal activation of the *GAL1* promoter (11), further demonstrating that the observed defects are specific for the Thr4 mutation. In all cases, reduced transcription of genes in T4V correlated with less RNAP II recruited to the genes' promoters, as determined by ChIP analysis (SI Appendix, Fig. S1D). The *htz1*Δ strain also grew normally on amino acid-depleted and high osmolarity media, but, although Htz1 is known to play a role in activation of GAL1 (39, 41, 42), our *htz1*Δ strain (and several additional clones derived from the W303 parental strain) failed to display a significant growth defect on galactose (Fig. 4A) (see discussion). Taken together, our results indicate that Thr4 is required for full induction of a subset of inducible genes and suggest that the role of Thr4 in induction is related to Htz1 function.

Eviction of Htz1 from promoter-associated nucleosomes is thought to stimulate gene activation by exposing promoter DNA (35). To address whether impaired activation of *PHO5*, *GAL1*, and *GAL7* in T4V cells reflected defective eviction of Htz1 from their respective promoters, we used strains expressing Htz1 as a 6x-HA epitope-tagged derivative (SI Appendix, Fig. S5A) and performed ChIP experiments with an HA antibody. Fig. 4C shows that, in WT cells, induction of *PHO5*, *PHO84*, *GAL1*, and *GAL7* resulted in a significant loss of Htz1 from their respective promoter regions, confirming that Htz1 eviction corresponds with activation of these genes. Consistent with our gene-expression analysis, the T4V mutation effectively abolished Htz1 loss from the *PHO5*, *GAL1*, and *GAL7* promoters during their induction, but not from *PHO84* (Fig. 4C). Defective Htz1 eviction on these genes is specific to the CTD Thr4 mutation because altering the CTD to all consensus repeats (CONS) did not affect Htz1 loss on the *GAL1* promoter (SI Appendix, Fig. S5B). At the uninduced *PHO5* and *PHO84* promoters, less Htz1 was detected in T4V cells than in WT (Fig. 4C), probably caused by nucleosome loss coinciding with derepression of these genes due to the Thr4 mutation (see discussion). Despite the lower levels of Htz1 in the T4V mutant, significant Htz1 eviction was observed for PHO84, but not for PHO5. These observations strongly link Htz1 promoter dynamics to Thr4 and indicate that Thr4 is required for Htz1 eviction from some inducible genes, thereby enabling their full induction.

INO80c Functions in Thr4-Mediated Stimulation of Htz1 Eviction

We next examined whether Thr4 stimulates Htz1 loss from specific activated promoters through an INO80c-dependent mechanism. Indeed, Htz1 ChIP analysis showed that Htz1 loss from promoters of *PHO5*, *GAL1*, and *GAL7* during induction was

defective in cells expressing a nonfunctional, truncated version of the Ino80 ATPase subunit of the complex (*ino80Δ900*) (Fig. 4C). Consistent with previous observations that Ino80 plays a general role in regulating Htz1 occupancy on the genome (36), *ino80Δ900* cells showed reduced Htz1 occupancy on both *PHO5* and *PHO84* promoters in normal growth conditions (Fig. 4C). During induction of these genes, however, loss of Ino80 resulted in impaired Htz1 eviction for *PHO5* but not *PHO84*, paralleling the results seen with the T4V mutation (Fig. 4C). In all cases, impaired Htz1 loss from promoters resulted in reduced induction of the respective genes (Fig. 5A). These observations, therefore, implicate both Thr4 and Ino80 in clearing Htz1 from a specific set of induced promoters and suggest that Ino80 mediates the effects of Thr4 on induction of these genes (i.e., genes showing a requirement for Thr4 for induction are genes that show Ino80-dependence for Htz1 eviction). In support of this notion, in addition to the genetic interaction we detected between T4V and *IES2* in the SGA screen (Fig. 3A), we detected a negative synthetic genetic interaction between *INO80* and the T4V mutation (Fig. 5B). Furthermore, like T4V, cells lacking Ino80 were unable to grow on low Pi medium, were highly sensitive to galactose, but were essentially unaffected by amino acid depletion or high osmolarity (Fig. 5C). Therefore, our data implicate Ino80 in Thr4-stimulated loss of Htz1 from *PHO5*, *GAL1*, and *GAL7* promoters, pointing to an unexpected role for the CTD in chromatin remodeling.

Discussion

Here, we have provided evidence that Thr4 of the CTD plays a critical but selective role in transcription in budding yeast. Like our previous analysis of Thr4 function in chicken cells (8), we found that this highly conserved CTD residue does not

play a general role in transcription, but functions in expression of specific genes. Thr4 is required for both repression of PHO genes and for full activation of a subset of inducible genes, including *PHO5*, *GAL1*, and *GAL7*. Although it is likely that Thr4 maintains PHO gene repression by regulating chromatin structure, we found that its role in activation involves stimulating the eviction of Htz1 specifically from promoters of the Thr4-sensitive genes. To carry out this function, we provided evidence that Thr4 recruits or stimulates INO80c, whose chromatin remodeling activity acts to remove Htz1 from those promoters. Below, we discuss these previously unidentified functions of the CTD.

Our data implicate Thr4 in both repression and activation of specific genes, and chromatin remodeling appears to be at play in both cases. Repression of PHO-regulated genes is accomplished by blocking access of the basic helix–loop–helix transcription factor Pho4 to its recognition elements upstream of the PHO genes, by keeping it out of the nucleus and through interference with positioned nucleosomes on its binding sites (43). Consequently, disruption of chromatin structure at PHO genes leads to derepression. For example, in the absence of Set1, loss of H3K4 methylation causes reduced nucleosome occupancy around the *PHO5* promoter, and, as a result, *PHO5* becomes derepressed (44). Under repressive conditions, Cbf1, another basic helix–loop–helix protein, binds chromatin at many of the same sites that become bound by Pho4 under conditions of low Pi (45). Cbf1 is required for normal nucleosomal positioning (46), and, in its absence, PHO genes become derepressed, even under normal Pi conditions, when most Pho4 is cytoplasmic (43). Interestingly, there is a high degree of overlap between genes that we found are up-regulated in T4V with genes up-regulated in *CBF1*-deleted cells (43), suggesting that a similar defect in nucleosomal

arrangement occurs in T4V cells. Indeed, the reduced levels of Htz1 at the *PHO5* and *PHO84* promoters that we detected in T4V cells in noninducing conditions likely reflect reduced nucleosome occupancy at these derepressed genes. It is likely, then, that Thr4 contributes to maintaining the nucleosomal structure necessary for repression of PHO genes.

Induction of *PHO5* and *GAL1*, two of the Thr4-sensitive genes we identified, has been studied in some detail, and Htz1 is known to play a role in this process. Although Htz1 is evicted from the *GAL1* promoter upon induction (39, 41) and is quickly redeposited upon returning to a repressed state (47), its effects are on *GAL1* activation, not repression. Cells lacking *HTZ1* show reduced recruitment of TBP, RNAP II, Mediator, SAGA, and Swi/Snf, with reduced *GAL1* expression as a consequence (39, 41, 42, 47). The effect is transient (47), however, and, in some cases, RNAP II is efficiently recruited to the induced *GAL1* promoter, and *GAL1* expression occurs in the absence of Htz1 (our results and refs. 39 and 48), indicating that other pathways can sometimes compensate for the lack of Htz1. Htz1 eviction also occurs at the *PHO5* promoter upon its induction (39, 49). Nucleosome disassembly at the *PHO5* promoter was shown to be rate-limiting for its expression (50), and efficient *PHO5* induction is dependent on Htz1 (our results and ref. 39). We have now shown that Thr4 is required for eviction of Htz1 from *GAL1* and *PHO5*, as well as *GAL7* promoters. The consequences of the T4V mutation on expression of these genes and growth on low Pi and galactose-containing media are more severe than in *htz1*Δ cells, indicating that, beyond promoting Htz1 eviction, Thr4 likely stimulates expression through additional pathways. Nonetheless, our data, supported by genetic interactions between T4V and

HTZ1 and its regulators, connect transcription to chromatin remodeling by stimulating Thr4-dependent Htz1 eviction.

Our data also implicate INO80c in Thr4 function. In addition to roles in double-strand break repair and stalled replication fork recovery (51), INO80c is required for full induction of a number of genes, including reporter genes containing upstream activation sequences of the *PHO5* or *GAL1* promoters (52, 53). INO80c associates with many RNAP II promoters and is recruited to genes when they are activated (54, 55), including *PHO5*, where it was shown to be required for nucleosome remodeling and gene induction (52). In the absence of Ino80, Htz1 localization across the genome is perturbed, with increased occupancy at some loci and reduced occupancy at others, thereby linking INO80c with Htz1 dynamics (36). More specifically, Ino80 is required for Htz1 eviction from the *KAR4* promoter during induction, demonstrating that it can function in clearing promoter-proximal nucleosomes of Htz1 coincident with gene activation (36). Indeed, we have now shown that additional genes, specifically those that are sensitive to the T4V mutation (i.e., *PHO5*, *GAL1*, and *GAL7*), also require Ino80 for efficient Htz1 loss during their activation. Based on our findings, we expect that other genes found to be dependent on Ino80 for Htz1 eviction will also show Thr4 sensitivity.

We have presented several lines of evidence linking Thr4 with Ino80 function. Based on these connections, we propose a model (Fig. 6) in which, upon activation, a recruited molecule of RNAP II triggers eviction of Htz1 from promoter-proximal nucleosomes by stimulating the INO80 complex, or facilitating its recruitment, to newly activated promoters, in a manner dependent on Thr4 or its phosphorylation. We were unable to detect an Ino80–Rpb1 interaction by coimmunoprecipitation, suggesting that

recruitment or stimulation of INO80c by Thr4 is transient and/or mediated by other factors. However, failure to detect such an interaction is not surprising as many known CTD interactions are not detectable in this type of assay. For example, capping enzymes, polyadenylation factors, and the CTD kinases Ctk1 and Kin28 were not detected in a tandem-affinity purification of Rpb1 even though they are all well-characterized CTD binding proteins (56). Importantly, transcription and RNAP II have been shown to be important for promoter-proximal nucleosome eviction (57). We now implicate Thr4 and INO80c in this process, which we propose facilitates further recruitment of additional RNAP II molecules. In vitro studies showed that INO80c evicts Htz1/H2B dimers from nucleosomal octamers and that, in some cases, eviction was part of an INO80c-dependent Htz1/H2B exchange for H2A/H2B (36) whereas, in another in vitro study, INO80c was not able to replace Htz1/H2B dimers with H2A/H2B dimers in nucleosomal arrays (58). In our model, however, we suggest that Thr4-mediated recruitment or stimulation of INO80c at specific promoters causes a loss of Htz1 that destabilizes promoter-proximal nucleosomes, thereby increasing accessibility to promoter DNA and facilitating further rounds of RNAP II recruitment and increased transcription (33, 35). Consequently, H2A-containing nucleosomes (which presumably form at promoters in the absence of Htz1) are not efficiently destabilized by INO80c, resulting in reduced gene activation.

Defects in repression or activation that result from the T4V mutant CTD might arise from an inability to phosphorylate the mutated Thr residues. Genomic ChIP analysis determined that the average Thr4 phosphorylation pattern in human cells increases toward the 3’ end of genes (24). In yeast, a genomic ChIP (59), as well as our

ChIP analysis of individual genes, showed that most Thr4 phosphorylation is detected on Rpb1 associated with the middle portion of genes, with only low levels detected at promoters. This distribution might indicate that the Thr residue itself is important for promoter-specific functions, including those we report in this study.

A recent study examined the transcriptomes of fission yeast strains harboring a variety of CTD substitutions (25). Consistent with our findings, mutation of Thr4 (to Ala; T4A) in that study affected expression of only a small number of genes, including orthologs of *PHO5* and *PHO84*. However, whereas our T4V mutation resulted in derepression of these genes, in fission yeast expressing the T4A mutant, expression of the orthologous genes was significantly decreased. The different effect of Thr4 mutation might reflect differences between the two species. For example, Ser2 has been shown to be required for viability in *S. cerevisiae* but not in *Schizosaccharomyces pombe* during normal growth (26, 60). Differences between the CTD constructs could also contribute (e.g., T4A and 14 consensus heptads in ref. 29, compared with T4V and 24 heptads in our study). Nonetheless, the fact that *PHO5* and *PHO84* (and orthologs) are among only a small number of genes whose expression relies on Thr4 in both budding and fission yeast strains strongly points to conserved roles for this CTD residue in regulating transcription of PHO genes.

An interesting question is whether our findings on Thr4 function extend to metazoan organisms. As mentioned above, two studies have characterized the role of Thr4 and its phosphorylation in vertebrate species, in which it plays functions necessary for viability (8, 24). Furthermore, INO80c functions in promoting eviction of transcription start site-associated nucleosomes containing the mammalian Htz1 ortholog H2A.Z

during gene activation in mouse embryonic stem cells during differentiation, indicating that the same basic INO80c-mediated mechanism is at play throughout eukaryotes (61). It remains to be seen whether the role we uncovered for Thr4 in mediating this mechanism in yeast is conserved in higher species. In any event, our results show that Thr4 of the yeast CTD plays an important role in coupling transcription with promoter-associated nucleosomal remodeling, providing yet another function for the multifunctional CTD.

Materials and Methods

A detailed description of yeast strain construction and growth conditions, ChIP, RNA and protein analyses, and RNA sequencing can be found in SI Appendix, SI Materials and Methods. Briefly, T4V and CONS strains were constructed by homologous recombination in yeast strains using CTD constructs prepared for our previous study (8). SGA analysis, comparing CONS and T4V strain genetic interactions, was as previously described (62). RT-PCR analyses were performed with random-primed total RNA for cDNA synthesis whereas RNA-seq experiments used polyA+ RNA. Primer sequences used for cloning, ChIP, and RT-PCR are available upon request.

Acknowledgments

We thank Elizabeth Miller and Silvere Pagant (Columbia University) for assistance with the SGA screen. This work was supported by National Institutes of Health Grants R01GM097174 (to J.L.M.) and R01GM084089 (to B.T.).

References

1. Rando OJ, Winston F. Chromatin and transcription in yeast. Genetics. 2012; 190(2): 351–387.

2. Smolle M, Workman JL. Transcription-associated histone modifications and cryptic transcription. Biochim Biophys Acta. 2013; 1829(1): 84–97.

3. Corden JL. RNA polymerase II C-terminal domain: Tethering transcription to transcript and template. Chem Rev. 2013; 113(11): 8423–8455.

4. Heidemann M, Hintermair C, Voß K, Eick D. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. Biochim Biophys Acta. 2013; 1829(1): 55–62.

5. Hsin JP, Manley JL. The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. 2012; 26(19): 2119–2137.

6. Drogat J, Hermand D. Gene-specific requirement of RNA polymerase II CTD phosphorylation. Mol Microbiol. 2012; 84(6): 995–1004.

7. Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S. Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. Science. 2007; 318(5857): 1777–1779.

8. Hsin JP, Sheth A, Manley JL. RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3’ end processing. Science. 2011; 334(6056): 683–686.

9. Tsai KL, Sato S, Tomomori-Sato C, Conaway RC, Conaway JW, Asturias FJ. A conserved Mediator-CDK8 kinase module association regulates Mediator-RNA polymerase II interaction. Nat Struct Mol Biol. 2013; 20(5): 611–619.

10. Myers LC, Gustafsson CM, Bushnell DA, Lui M, Erdjument-Bromage H, Tempst P, Kornberg RD. The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain. Genes Dev. 1998; 12(1): 45–54.

11. Allison LA, Ingles CJ. Mutations in RNA polymerase II enhance or suppress mutations in GAL4. Proc Natl Acad Sci USA. 1989; 86(8): 2794–2798.

12. Gerber HP, Hagmann M, Seipel K, Georgiev O, West MAL, Litingtung Y, Schaffner W, Corden JL. RNA polymerase II C-terminal domain required for enhancer-driven transcription. Nature. 1995; 374(6523): 660–662.

13. Scafe C, Chao D, Lopes J, Hirsch JP, Henry S, Young RA. RNA polymerase II C-terminal repeat influences response to transcriptional enhancer signals. Nature. 1990; 347(6292): 491–494.

14. Nagaike T, Manley JL. Transcriptional activators enhance polyadenylation of mRNA precursors. RNA Biol. 2011; 8(6): 964–967.

15. Nogues G, Kadener S, Cramer P, Bentley D, Kornblihtt AR. Transcriptional activators differ in their abilities to control alternative splicing. J Biol Chem. 2002; 277(45): 43110–43114.

16. Rosonina E, Bakowski MA, McCracken S, Blencowe BJ. Transcriptional activators control splicing and 3'-end cleavage levels. J Biol Chem. 2003; 278(44): 43034–43040.

17. Bataille AR, Jeronimo C, Jacques PE, Laramee L, Fortin ME, Forest A, Bergeron M, Hanes SD, Robert F. A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. Mol Cell. 2012; 45(2): 158–170.

18. Kim H, Erickson B, Luo W, Seward D, Graber JH, Pollock D, Megee PC, Bentley DL. Gene-specific RNA polymerase II phosphorylation and the CTD code. Nat Struct Mol Biol. 2010; 17(10): 1279–1286.

19. McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, Hessel A, Foster S, Shuman S, Bentley DL. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes Dev. 1997; 11(24): 3306–3318.

20. Cho EJ, Takagi T, Moore CR, Buratowski S. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. Genes Dev. 1997; 11(24): 3319–3326.

21. Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M, Greenblatt JF, Shilatifard A. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: Linking transcriptional elongation to histone methylation. Mol Cell. 2003; 11(3): 721–729.

22. Ng HH, Robert F, Young RA, Struhl K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol Cell. 2003; 11(3): 709–719.

23. Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol. 2005; 25(8): 3305–3316.

24. Hintermair C, Heidemann M, Koch F, Descostes N, Gut M, Gut I, Fenouil R, Ferrier P, Flatley A, Kremmer E, Chapman RD, Andrau JC, Eick D. Threonine-4 of mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. EMBO J. 2012; 31(12): 2784–2797.

25. Schwer B, Bitton DA, Sanchez AM, Bahler J, Shuman S. Individual letters of the RNA polymerase II CTD code govern distinct gene expression programs in fission yeast. Proc Natl Acad Sci USA. 2014; 111(11): 4185–4190.

26. Schwer B, Shuman S. Deciphering the RNA polymerase II CTD code in fission yeast. Mol Cell. 2011; 43(2): 311–318.

27. Allison LA, Moyle M, Shales M, Ingles CJ. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. Cell. 1985; 42(2): 599–610.

28. Ogawa N, DeRisi J, Brown PO. New components of a system for phosphate accumulation and polyphosphate metabolism in Saccharomyces cerevisiae revealed by genomic expression analysis. Mol Biol Cell. 2000; 11(12): 4309–4321.

29. Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Page N, Robinson M, Raghibizadeh S, Hogue CW, Bussey H, Andrews B, Tyers M, Boone C. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science. 2001; 294(5550): 2364–2368.

30. Kobor MS, Venkatasubrahmanyam S, Meneghini MD, Gin JW, Jennings JL, Link AJ, Madhani HD, Rine J. A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. PLoS Biol. 2004; 2(5): E131.

31. Krogan NJ, Keogh MC, Datta N, Sawa C, Ryan OW, Ding H, Haw RA, Pootoolal J, Tong A, Canadien V, Richards DP, Wu X, Emili A, Hughes TR, Buratowski S, Greenblatt JF. A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol Cell. 2003; 12(6): 1565–1576.

32. Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science. 2004; 303(5656): 343–348.

33. Li B, Pattenden SG, Lee D, Gutierrez J, Chen J, Seidel C, Gerton J, Workman JL. Preferential occupancy of histone variant H2AZ at inactive promoters influences local histone modifications and chromatin remodeling. Proc Natl Acad Sci USA. 2005; 102(51): 18385–18390.

34. Venters BJ, Pugh BF. A canonical promoter organization of the transcription machinery and its regulators in the Saccharomyces genome. Genome Res. 2009; 19(3): 360–371.

35. Zhang H, Roberts DN, Cairns BR. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. Cell. 2005; 123(2): 219–231.

36. Papamichos-Chronakis M, Watanabe S, Rando OJ, Peterson CL. Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. Cell. 2011; 144(2): 200–213.

37. Raisner RM, Hartley PD, Meneghini MD, Bao MZ, Liu CL, Schreiber SL, Rando OJ, Madhani HD. Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. Cell. 2005; 123(2): 233–248.

38. Guillemette B, Bataille AR, Gevry N, Adam M, Blanchette M, Robert F, Gaudreau L. Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. PLoS Biol. 2005; 3(12): e384.

39. Santisteban MS, Kalashnikova T, Smith MM. Histone H2A.Z regulats transcription and is partially redundant with nucleosome remodeling complexes. Cell. 2000; 103(3): 411–422.

40. Oshima Y. The phosphatase system in Saccharomyces cerevisiae. Genes Genet Syst. 1997; 72(6): 323–334.

41. Adam M, Robert F, Larochelle M, Gaudreau L. H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. Mol Cell Biol. 2001; 21(18): 6270–6279.

42. Halley JE, Kaplan T, Wang AY, Kobor MS, Rine J. Roles for H2A.Z and its acetylation in GAL1 transcription and gene induction, but not GAL1-transcriptional memory. PLoS Biol. 2010; 8(6): e1000401.

43. Zhou X, O'Shea EK. Integrated approaches reveal determinants of genome-wide binding and function of the transcription factor Pho4. Mol Cell. 2011; 42(6): 826–836.

44. Wang SS, Zhou BO, Zhou JQ. Histone H3 lysine 4 hypermethylation prevents aberrant nucleosome remodeling at the PHO5 promoter. Mol Cell Biol. 2011; 31(15): 3171–3181.

45. Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, Hannett NM, Tagne JB, Reynolds DB, Yoo J, Jennings EG, Zeitlinger J, Pokholok DK, Kellis M, Rolfe PA, Takusagawa KT, Lander ES, Gifford DK, Fraenkel E, Young RA. Transcriptional regulatory code of a eukaryotic genome. Nature. 2004; 431(7004): 99–104.

46. Kent NA, Eibert SM, Mellor J. Cbf1p is required for chromatin remodeling at promoter-proximal CACGTG motifs in yeast. J Biol Chem. 2004; 279(26): 27116–27123.

47. Gligoris T, Thireos G, Tzamarias D. The Tup1 corepressor directs Htz1 deposition at a specific promoter nucleosome marking the GAL1 gene for rapid activation. Mol Cell Biol. 2007; 27(11): 4198–4205.

48. Santisteban MS, Hang M, Smith MM. Histone variant H2A.Z and RNA polymerase II transcription elongation. Mol Cell Biol. 2011; 31(9): 1848–1860.

49. Boeger H, Griesenbeck J, Strattan JS, Kornberg RD. Removal of promoter nucleosomes by disassembly rather than sliding in vivo. Mol Cell. 2004; 14(5): 667–673.

50. Boeger H, Griesenbeck J, Kornberg RD. Nucleosome retention and the stochastic nature of promoter chromatin remodeling for transcription. Cell. 2008; 133(4): 716–726.

51. Bao Y, Shen X. INO80 subfamily of chromatin remodeling complexes. Mutat Res. 2007; 618(1-2): 18–29.

52. Barbaric S, Luckenback T, Schmid A, Blaschke D, Horz W, Korber P. Redundancy of chromatin remodeling pathways for the induction of the yeast PHO5 promoter in vivo. J Biol Chem. 2007; 282(38): 27610–27621.

53. Ebbert R, Birkmann A, Schüller HJ. The product of the SNF2/SWI2 paralogue INO80 of Saccharomyces cerevisiae required for efficient expression of various yeast structural genes is part of a high-molecular-weight protein complex. Mol Microbiol. 1999; 32(4): 741–751.

54. Jónsson ZO, Jha S, Wohlschlegel JA, Dutta A. Rvb1p/Rvb2p recruit Arp5p and assemble a functional Ino80 chromatin remodeling complex. Mol Cell. 2004; 16(3): 465–477.

55. Klopf E, Paskova L, Sole C, Mas G, Petryshyn A, Posas F, Wintersberger U, Ammerer G, Schuller C. Cooperation between the INO80 complex and histone chaperones determines adaptation of stress gene transcription in the yeast Saccharomyces cerevisiae. Mol Cell Biol. 2009; 29(18): 4994–5007.

56. Krogan NJ, et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature. 2006; 440(7084): 637–643.

57. Weiner A, Hughes A, Yassour M, Rando OJ, Friedman N. High-resolution nucleosome mapping reveals transcription-dependent promoter packaging. Genome Res. 2010; 20(1): 90–100.

58. Luk E, Ranjan A, Fitzgerald PC, Mizuguchi G, Huang Y, Wei D, Wu C. Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. Cell. 2010; 143(5): 725–736.

59. Mayer A, Heidemann M, Lidschreiber M, Schreieck A, Sun M, Hintermair C, Kremmer E, Eick D, Cramer P. CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. Science. 2012; 336(6089): 1723–1725.

60. West ML, Corden JL. Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. Genetics. 1995; 140(4): 1223–1233.

61. Li Z, Gadue P, Chen K, Jiao Y, Tuteja G, Schug J, Li W, Kaestner KH. Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. Cell. 2012; 151(7): 1608–1616.

62. Tong AH, Boone C. Synthetic genetic array analysis in Saccharomyces cerevisiae. Methods Mol Biol. 2006; 313: 171–192.

63. Rosonina E, Duncan SM, Manley JL. SUMO functions in constitutive transcription and during activation of inducible genes in yeast. Genes Dev. 2010; 24(12): 1242–1252.

Figure 1

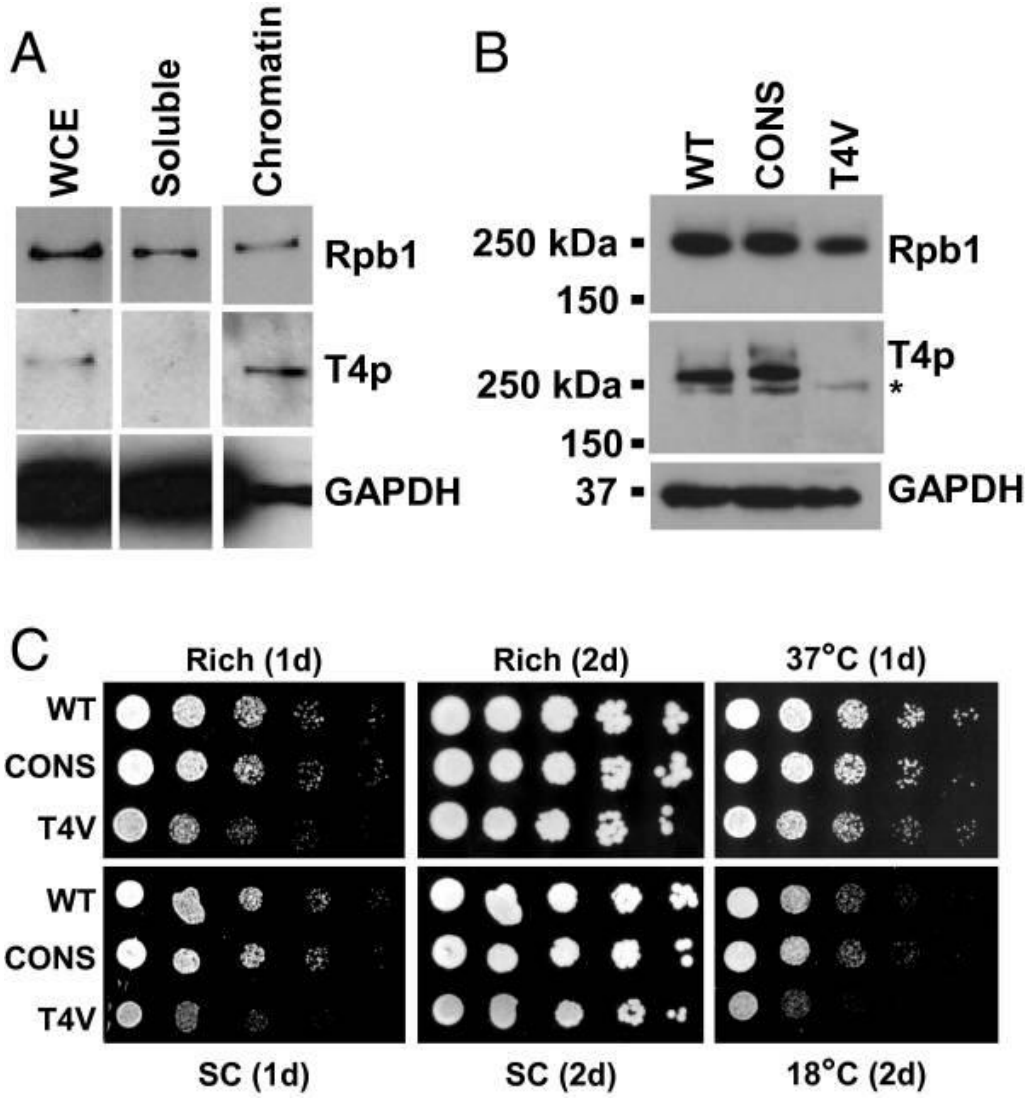


Figure 1: Thr4 of the yeast CTD is phosphorylated on chromatin-associated RNAP II and is not essential for viability. (A) Chromatin fractionation of yeast extract followed by Western blot analysis with an RNAP II antibody (Rpb1), an antibody that recognizes phosphorylated Thr4 on CTD repeats (T4p) (8), and a GAPDH antibody. (B) Western blot analysis of yeast extracts derived from WT, CONS, or T4V strains analyzed with the same antibodies as in A. The asterisk (*) indicates a nonspecific cross-reacting band detected at longer exposures. (C) Yeast spot assay comparing growth of WT, CONS, and T4V strains, with serial fivefold dilutions, on rich or synthetic complete (SC) medium at 30 ° C, or on SC medium at indicated temperature, for indicated number of days.

Figure 2

A

GENE	T4V vs WT Log2	T4V vs WT P-value	CONS vs WT Log2	CONS vs WT P-value
UPREGULATED				
*PHO84	10.8	**	1.68	2.3E-02
*PHO89	6.6	5.2E-107	NA	NA
*SPL2	6.2	1.6E-22	0.29	1.0E+00
*PHO5	3.9	1.4E-250	-0.13	1.0E+00
*PHO11	3.5	5.6E-15	NA	NA
*PHO12	3.1	1.3E-26	0.35	4.0E-01
GIT1	2.7	2.2E-32	-0.45	5.1E-01
YIL024C	2.5	1.5E-23	NA	NA
SNZ1	2.3	1.0E-63	0.10	6.2E-01
SNR62	2.3	1.1E-10	NA	NA
*VTC3	2.1	**	0.36	2.1E-01
YLR312C	2.1	2.6E-15	NA	NA
SUL2	2.1	4.1E-11	-0.20	7.2E-01
ADE17	1.9	2.6E-141	-0.02	1.0E+00
DAL80	1.9	4.9E-14	1.20	1.6E-01
*VTC4	1.7	4.6E-237	-0.27	2.7E-02
MEP2	1.7	3.5E-104	0.57	1.5E-17
DOWNREGULATED				
SPO22	-3.1	1.5E-16	0.40	3.5E-01
LOT5	-2.8	2.1E-25	-0.16	3.0E-01
SNR86	-2.3	3.7E-22	NA	NA
GCD10	-2.2	1.7E-49	-0.12	1.8E-01
PTC2	-2.1	1.6E-200	-0.40	4.7E-08
AIM4	-2.1	2.7E-26	0.03	1.0E+00
HMS2	-2.0	2.0E-24	-1.14	8.1E+05
MOD5	-2.0	1.6E-19	0.27	4.7E-02
SEC39	-1.6	5.7E-16	0.35	9.1E-02

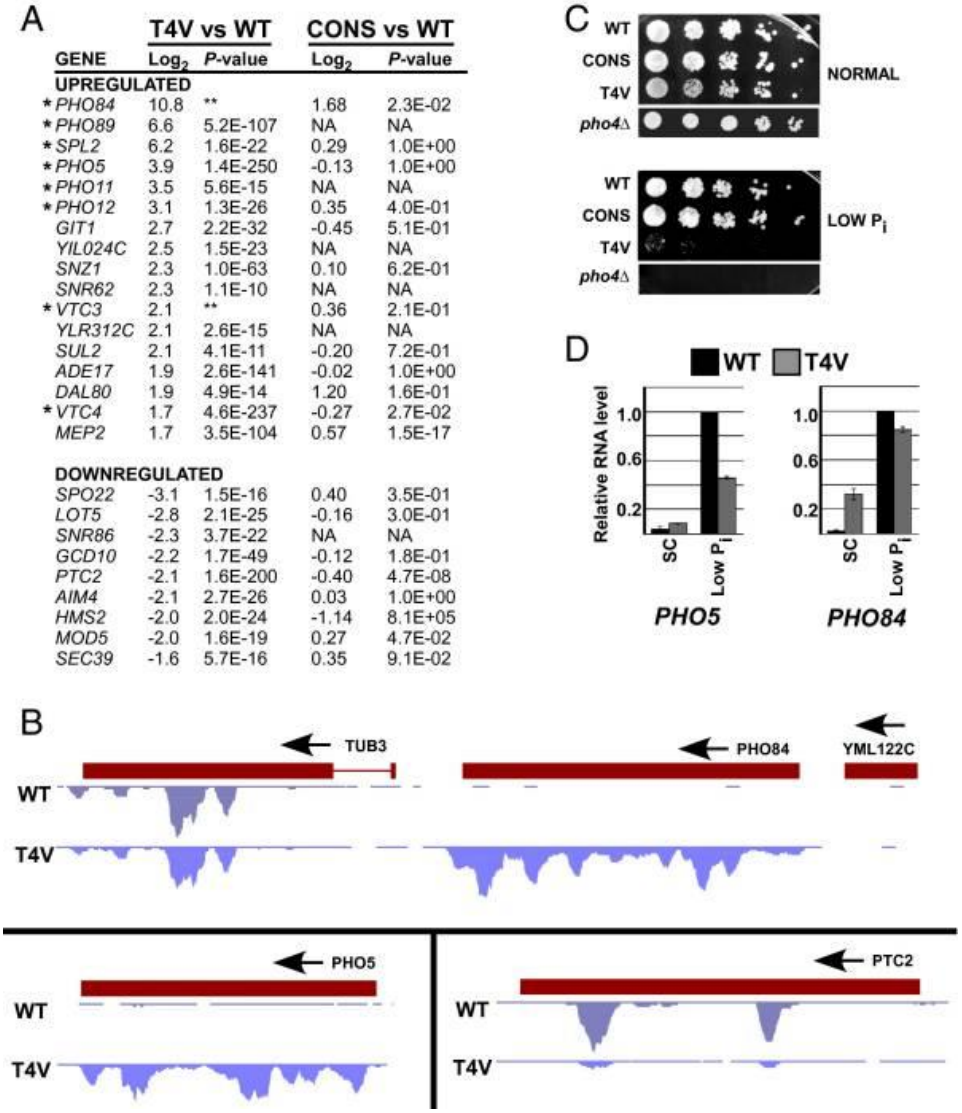


Figure 2: Thr4 is required to maintain repression of PHO genes under normal growth conditions. (A) Summary of most significantly up- and down-regulated genes from RNA-seq analysis performed in WT and T4V strains grown in SC medium. Average change in expression of genes (log2) in T4V over WT strains and P value (Fisher's exact test) were determined from two sets of data and are indicated here (refer to Dataset S1 for a complete list and details). Corresponding log2 and P values from the WT versus CONS RNA-seq analysis (Dataset S2) are also shown. Genes previously demonstrated to be up-regulated during conditions of low Pi (28) are indicated (*). P values that are beyond the limit of our analysis are indicated with double asterisk (**). NA, genes whose expression was not detectable in the sample. (B) Read distributions from the WT versus CONS RNA-seq analysis for selected genes including PHO84 (with flanking genes shown), PHO5, and PTC2. (C) Spot assay comparing growth of indicated strains and serial fivefold dilutions on normal or low Pi medium. (D) RT-PCR analysis of PHO5 and PHO84 RNA levels in SC or low Pi media in WT or htz1Δ strains.

Figure 3

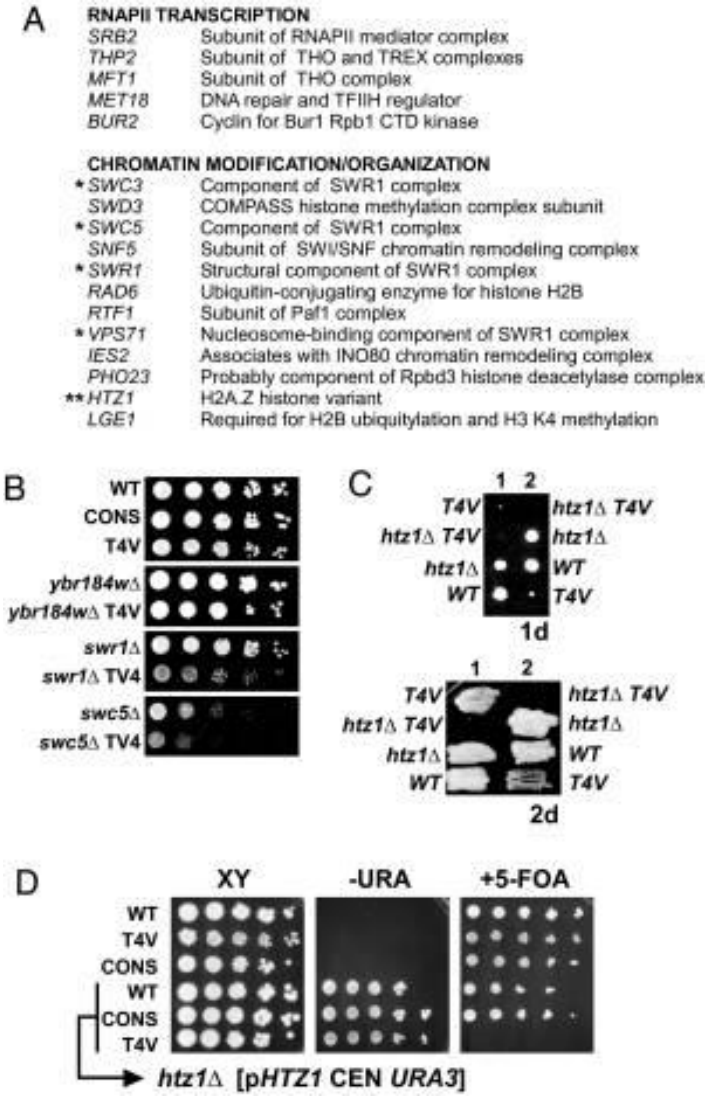


Figure 3: Thr4 is linked genetically to Htz1 and its regulators. (A) List of genes showing synthetic genetic interaction with T4V mutation in two SGA screens. Genes displaying strongest interactions and involved in RNAP II transcription or chromatin regulation are shown (refer to SI Appendix, Fig. S2 for a complete list). (B) Confirmation of genetic interaction between T4V and SWR1 and SWC5 components of SWR1 complex. Strains of indicated genotypes were generated by recombinant transformation and growth compared by spot assay on SC medium with serial fivefold dilutions. YBR184W showed no interaction with T4V in SGA screen and serves as a control. (C) Tetrad dissection analyses demonstrating synthetic lethal phenotype between T4V and HTZ1. Growth after 24 h is shown. After 48 h, colonies were streaked onto rich-medium plates, and growth is shown after additional 48 h. Genotype of inviable double mutant was inferred based on expected segregation pattern. (D) WT, T4V, and CONS strains were transformed with plasmid (pHTZ1 CEN URA3), and then endogenous HTZ1 was deleted by homologous recombination. Original and product strains were spotted on media shown, including 5-fluoroorotic acid (5-FOA), which allows growth only in the absence of URA3 expression (i.e., if the plasmid is lost).

Figure 4

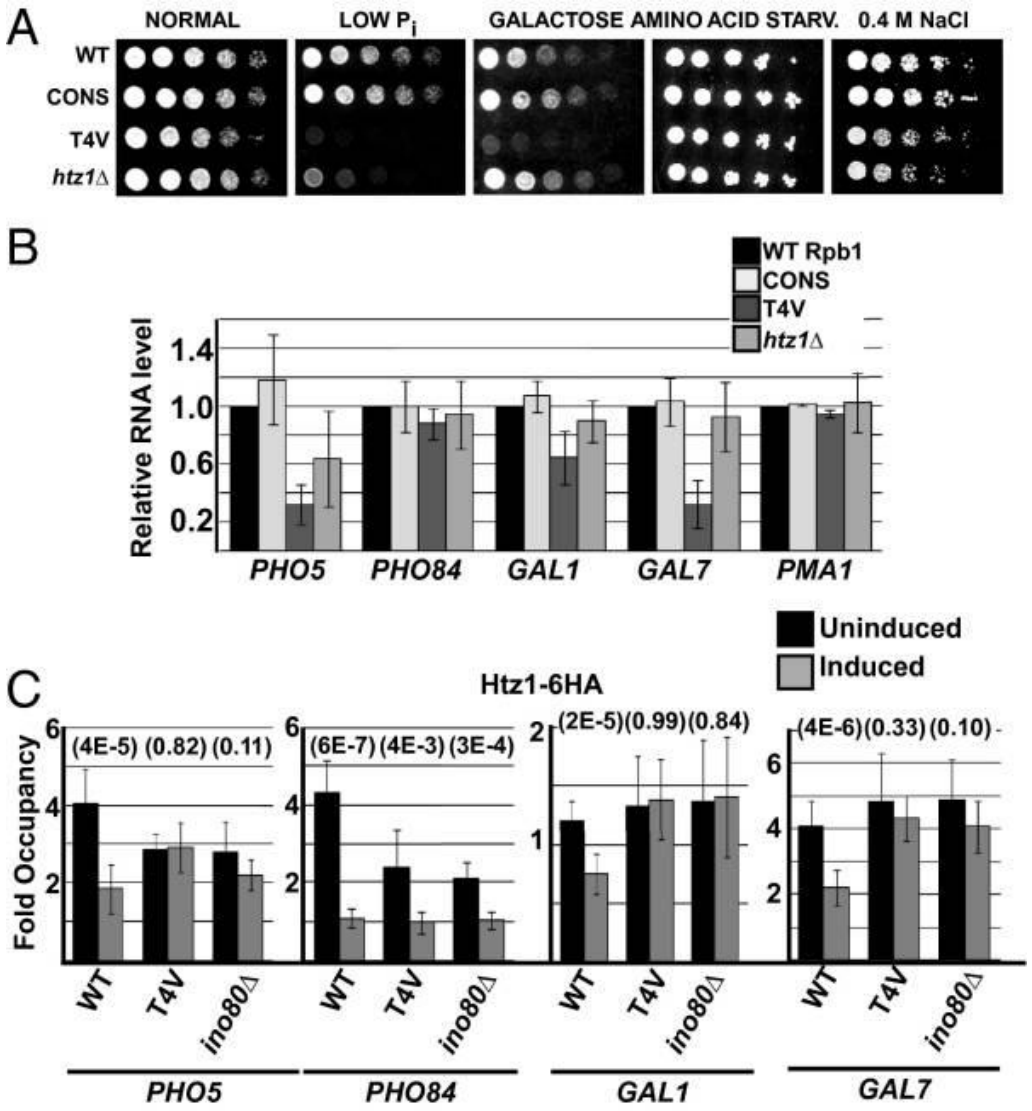


Figure 4: Thr4 is required for full expression and Htz1 eviction from promoters of specific genes. (A) Spot assay comparing growth of indicated strains and serial fivefold dilutions on indicated medium. Amino acid starvation medium was previously described (63). (B) RT-PCR analysis of mRNA levels of indicated genes under the respective inducing conditions, with PMA1 levels analyzed as control, grown in SC medium. Data are represented as mean ± SD of three independent experiments. (C) ChIP analysis of Htz1-6HA occupancy at promoter regions of indicated genes, uninduced or induced as indicated, in WT, T4V, and ino80Δ strains. Data are represented as mean ± SD of three independent experiments, and P values for each pair of values (uninduced and induced) are indicated in parentheses above bars.

Figure 5

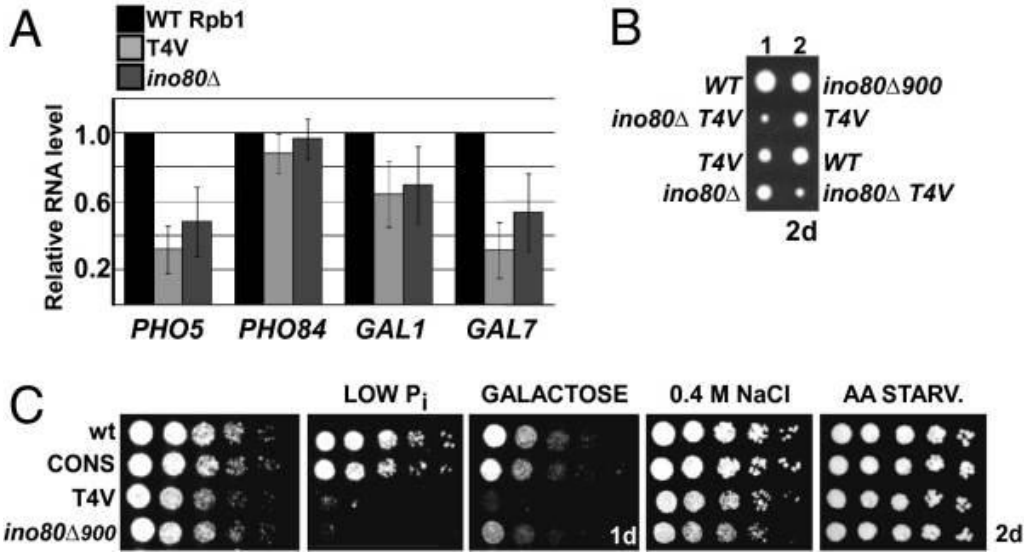


Figure 5: Evidence that Thr4 dependence is linked to the INO80 complex. (A) RT-PCR analysis of mRNA levels of indicated genes under respective induced conditions in indicated strains. Data are represented as mean ± SD of three independent experiments. (B) Tetrad dissection analyses demonstrating negative synthetic genetic interaction between T4V and INO80. Growth after 48 h is shown. (C) Growth comparison of indicated stains and serial fivefold dilutions in indicated media.

Figure 6

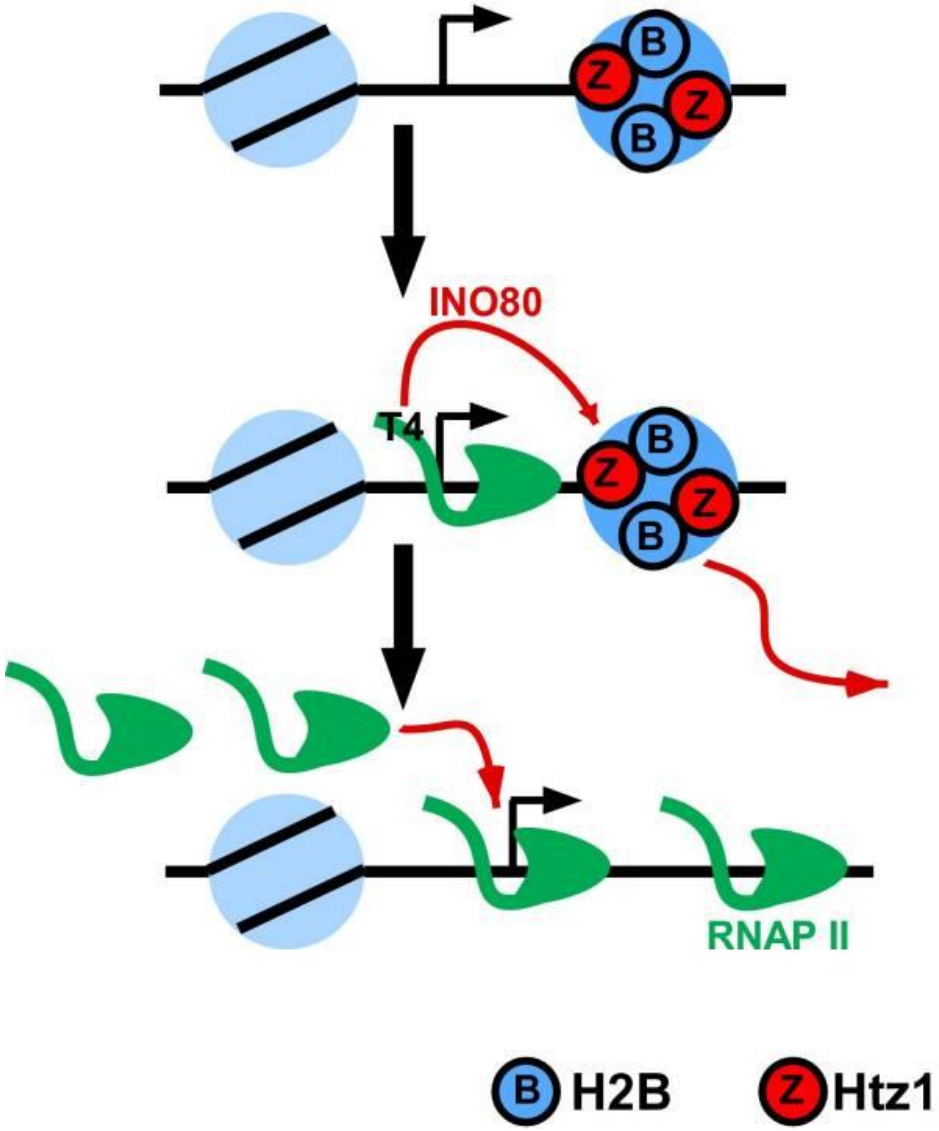


Figure 6: Model indicating role for Thr4 of yeast CTD repeats in promoting Htz1 eviction. On repressed genes, promoter-proximal nucleosomes contain Htz1/H2B dimers (depicted as encircled Z and B). During activation, an initially recruited RNAP II molecule stimulates INO80 activity or recruitment through Thr4 or Thr4 phosphorylation, thereby enabling INO80-mediated eviction of Htz1/H2B from nucleosomes. Htz1 eviction destabilizes nucleosomes, facilitating their loss and increasing DNA access for further rounds of transcription. When transcription ends, in the absence of Thr4, INO80 no longer acts, and Htz1/H2B is redeposited at promoter-proximal nucleosomes by the SWR1 complex.

SUPPLEMENTARY INFORMATION

Supplementary Materials and Methods

Supplementary Figure S1

Additional ChIP analyses supporting phosphorylation of Thr4 on RNAP II-transcribed genes

Supplementary Figure S2

RT-PCR analysis validation of RNA-seq data showing elevated expression of PHO gene RNAS in T4V strain

Supplementary Figure S3

Scatterplot comparing RNA-seq analyses of WT versus T4V and CONS strains

Supplementary Figure S4

Complete results for T4V SGA analysis.

Supplementary Figure S5

Analysis of Htz1 eviction in CONS strain

Supplementary Tables S1, S2 and S3

RNA seq statistics for Datasets S1 (Table S1) and S2 (Table S2); List of yeast strains used in this study (Table S3)

Datasets S1 and S2

Full RNA sequencing results for CONS vs T4V analysis (Dataset S1; RNA-seq) and CONS vs WT analysis (Dataset S2; 3′READS), as Excel spreadsheets

MATERIALS AND METHODS

Yeast strains and growth media.

Yeast strains used are listed in Table S3. T4V and CONS strains were constructed by PCR amplification of corresponding CTDs generated in preparation for our previous study (1), followed by transformation into haploid S288C or W303 strains. WT, CONS, and T4V strains were marked with NAT-MX cassette 3′ of the *RPO21* gene. Low P_i medium was prepared by dissolving 10 g/L yeast extract, 20 g/L peptone and 2.45 g/L magnesium sulfate in water, then adding drop-wise 8 mL/L ammonium hydroxide while mixing vigorously. The medium was allowed to sit for several hours on a bench to allow salts to precipitate, which were then filtered twice by vacuum filtration (0.22 μm). Glucose was then added to 2% (w/v), the pH was adjusted to ~7.0 with hydrochloric acid, and the medium was autoclaved. For solid medium, agar was added after pH adjustment. A yeast strain lacking *PHO4* was unable to grow specifically on the low P_i medium confirming it was successfully depleted of phosphate (Figure 2C). For liquid growth on low P_i medium, samples that had grown overnight in SC medium were washed then placed in low P_i medium for 5-6 hours before analysis. For liquid growth in galactose medium, samples were first grown in SC plus raffinose (2% w/v), then galactose was added to 5% (w/v) for one hour before samples were analyzed. SGA screen, yeast spot assays and media used were as previously described (2, 3).

ChIP, RNA and protein analyses.

Chromatin fractionation was performed as previously described (4) in strain W303a. For ChIP analyses, 50-mL cultures were grown and cells were lysed by bead-beating three times for one min each with one min in between in a cold room, but, otherwise, samples were processed and analyzed, and data quantified, as previously described (3). Western blot analyses and semiquantitative PCR were as previously described (2, 3, 5). Random hexamer oligonucleotides

were used with total RNA for the RT-PCR analyses. Where statistical analysis was performed for ChIP, a two-tailed Student's *t*-test was used, with *P*-values indicated above bars in the graph. Error bars in graphs represent standard deviations of at least three experiments. Antibodies used for ChIP and western blotting were HA (abm), GAPDH (Sigma), T4p (Novus), and Rpb1 (y-80; Santa Cruz). Primer sequences used for cloning, ChIP and RT-PCR analyses are available upon request.

RNA sequencing.

For RNA sequencing analyses, the integrity of total RNA was measured by Agilent Bioanalyzer. RNA samples with an integrity number >8.0 were used for further processing. For T4V vs WT samples, total RNA was subjected to 2 rounds of poly(A) selection using oligo-dT magnetic beads (NEB) followed by fragmentation using Ambion's RNA fragmentation kit at 70°C for 5 min. Fragmented RNA was purified by ethanol precipitation followed by dephosphorylation using recombinant shrimp alkaline phosphatase (NEB) and then phosphorylation using T4 kinase (NEB). Phosphorylated RNA was then purified by RNeasy kit (Qiagen) and was sequentially ligated to a 5′-adenylated 3′ adapter (5′/5rApp/TGGAATTCTCGGGTGCCAAGG/3ddC/) with the truncated T4 RNA ligase 2 (NEB) and to a 5′ adapter (5′-GUUCAGAGUUCUACAGUCCGACGAUC) with T4 RNA ligase 1 (NEB). The resultant RNA was reverse-transcribed to cDNA with Superscript III (Invitrogen), and the cDNA was amplified by 12 cycles of PCR with Phusion high fidelity polymerase (NEB). Amplified cDNAs were separated on 8% polyacrylamide gels. cDNAs containing an inset size ~100 nt were purified from the gel based on the size. The quality and quantity of the cDNA library were evaluated by Agilent Bioanalyzer and qPCR analysis. The strand-specific cDNA libraries generated by this protocol were sequenced on an Illumina Genome Analyzer IIx. RNA sequencing of CONS and WT samples in SC or XY media (Dataset S2) was performed using

3′READS, as previously described (6). Both RNA sequencing methods produced equivalent results, as determined by comparison of T4V vs WT samples analyzed by both methods.

REFERENCES

1. Hsin JP, Sheth A & Manley JL (2011) RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. *Science* 334(6056): 683-6.

2. Rosonina E, Duncan SM & Manley JL (2010) SUMO functions in constitutive transcription and during activation of inducible genes in yeast. *Genes Dev* 24(12): 1242-1252.

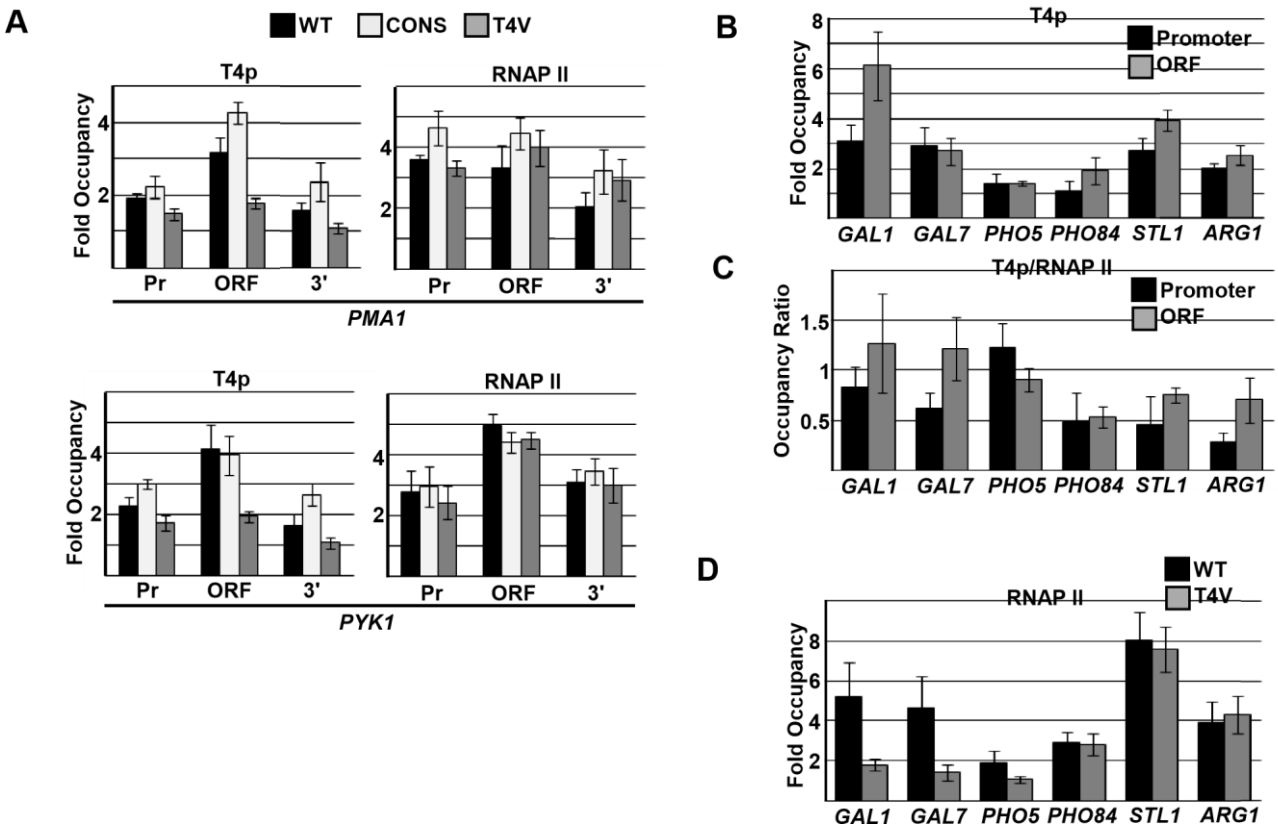
3. Rosonina E, Willis IM & Manley JL (2009) Sub1 functions in osmoregulation and in transcription by both RNA polymerases II and III. *Mol Cell Biol* 29(8): 2308-2321.

4. Svejstrup JQ, Petrakis TG & Fellows J (2003) Purification of elongating RNA polymerase II and other factors from yeast chromatin. *Methods Enzymol* 371: 491-8.

5. Rosonina E, Duncan SM & Manley JL (2012) Sumoylation of transcription factor Gcn4 facilitates its Srb10-mediated clearance from promoters in yeast. *Genes Dev* 26(4): 350-355.

6. Hoque M, *et al* (2013) Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods* 10(2): 133-139.

Supplementary Figure S1.



Thr4 is phosphorylated on RNAP II-transcribed genes.

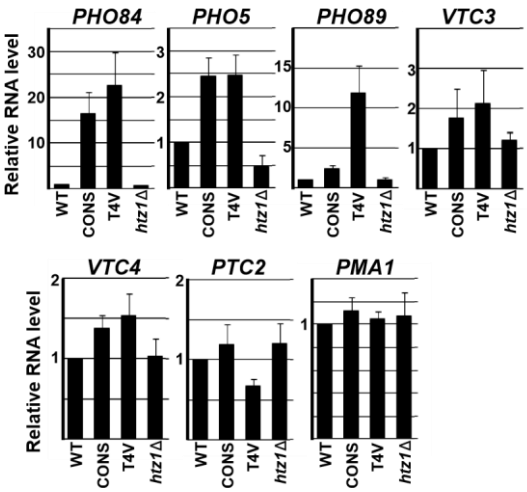
(A) ChIP analysis of phosphorylated Thr4 on CTD repeats (T4p) or RNAP II in WT, CONS, or T4V strains at promoter-proximal (Pr), ORF, or 3' regions of constitutively expressed *PMA1* and *PYK1* genes. Data are represented as mean +/- SD of three independent experiments.

(B) ChIP analysis of phosphorylated Thr4 on CTD repeats (T4p) on promoter-proximal or ORF regions of indicated genes under their respective induced conditions. Data are represented as mean +/- SD of three independent experiments.

(C) ChIP values from *B* were normalized to RNAP II levels, determined in the same samples at the same positions with an Rpb1 antibody. Data are represented as mean +/- SD of three independent experiments.

(D) ChIP analysis of RNAP II occupancy at promoters of indicated genes under their respective induced conditions in WT or T4V cells. Data are represented as mean +/- SD of three independent experiments.

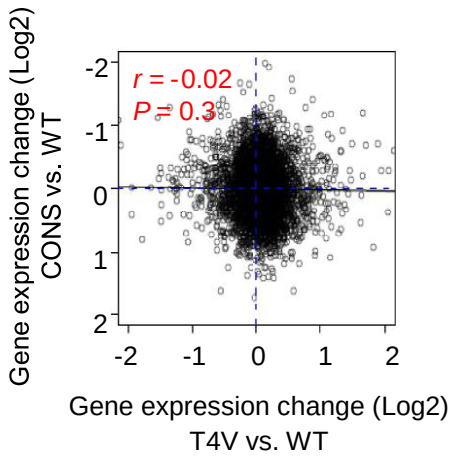
Supplementary Figure S2.



Elevated expression of PHO gene RNAs in T4V

RT-PCR validation of selected up- or down-regulated genes, with *PMA1* control, in WT, CONS, T4V, and *htz1*Δ strains. Total RNA from samples grown in normal (SC) medium was analyzed by semi-quantitative RT-PCR using random hexamer oligonucleotides for cDNA synthesis.

Supplementary Figure S3.



(4,610 genes with average read # >=20)

Comparison of RNAsequencing in CONS vs WT and T4V vs WT analyses.

Scatterplot comparing gene expression changes from CONS vs WT (Dataset S1) and T4V vs WT (Dataset S2) RNA sequencing analyses.

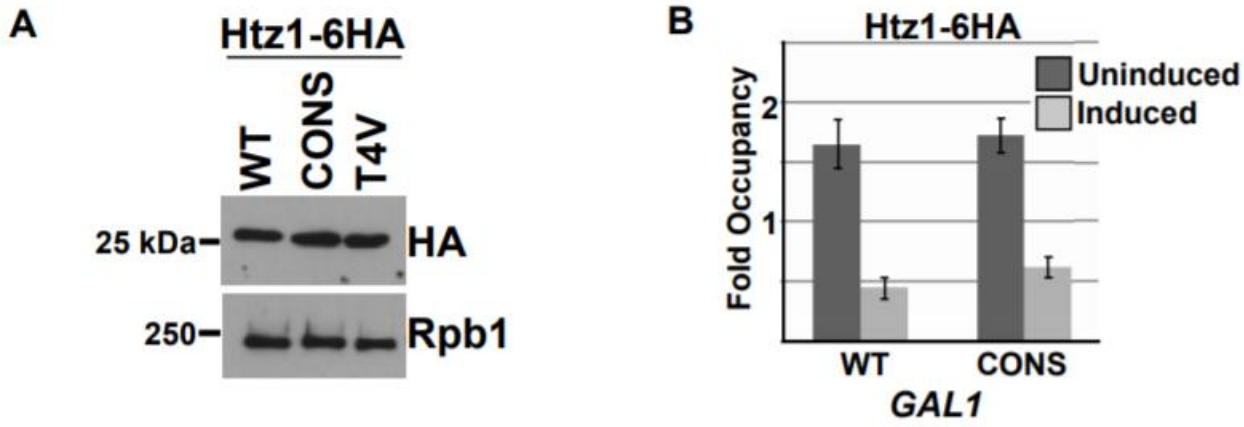
Supplementary Figure S4.

RNAP II TRANSCRIPTION			
	YHR041C	*SRB2*	Subunit of the RNA polymerase II mediator complex
	YHR167W	*THP2*	Subunit of the THO and TREX complexes
	YML062C	*MFT1*	Subunit of the THO complex
	YIL128W	*MET18*	DNA repair and TFIIH regulator
	YOL051W	*GAL11*	Subunit of the RNA polymerase II mediator complex
	YLR226W	*BUR2*	Cyclin for Bur1 Rpb1 CTD kinase
Chromatin Modification/Organization:			
*	YAL011W	*SWC3*	Component of the SWR1 complex
	YBR175W	*SWD3*	Subunit of the COMPASS (Set1C) histone methylation complex
*	YBR231C	*SWC5*	Component of the SWR1 complex
	YBR289W	*SNF5*	Subunit of the SWI/SNF chromatin remodeling complex
*	YDR334W	*SWR1*	Swi2/Snf2-related ATPase that is the structural component of the SWR1 complex
	YGL058W	*RAD6*	Ubiquitin-conjugating enzyme (E2); ubiquitylates histone H2B at K123
	YGL244W	*RTF1*	Subunit of RNAPII-associated chromatin remodeling Paf1 complex
*	YML041C	*VPS71*	Nucleosome-binding component of the SWR1 complex
	YNL215W	*IES2*	Protein that associates with the INO80 chromatin remodeling complex
	YNL097C	*PHO23*	Probable component of the Rpd3 histone deacetylase complex
**	YOL012C	*HTZ1*	Histone variant H2AZ, exchanged for histone H2A in nucleosomes by the SWR1 complex
	YPL055C	*LGE1*	Gene requried for H2B ubiquitination and H3 K4 methylation (REF: Hwang WW .. Madhani HD Mol Cell 2003)
OTHER			
	YBR275C	*RIF1*	Telomere maintenance
	YER122C	*GLO3*	ADP-ribosylation factor GTPase activating protein
	YER151C	*UBP3*	Ubiquitin-specific protease involved in transport and osmotic response; interacts with Bre5p
	YHR043C	*DOG2*	2-deoxyglucose-6-phosphate phosphatase
	YIL006W	*YIA6*	Mitochondrial NAD+ transporter
	YKL204W	*EAP1*	eIF4E-associated protein
	YMR198W	*CIK1*	Kinesin-associated protein required for both karyogamy and mitotic spindle organization
	YNL248C	*RPA49*	RNA polymerase I subunit A49
	YNR027W	*BUD17*	Putative pyridoxal kinase
	YNR047W	*FPK1*	Ser/Thr protein kinase that regulates the putative phospholipid translocases Lem3p-Dnf1p/Dnf2p
	YNR051C	*BRE5*	Ubiquitin protease cofactor, forms deubiquitination complex with Ubp3
	YPR101W	*SNT309*	Member of the NineTeen Complex (NTC) s

List of genes strongly interacting genetically with T4V mutation in SGA screens

Two SGA analyses were performed with WT and T4V strains. Stains expressing T4V and each of ~4700 deleted non-essential genes were inspected for growth defects compared to the same gene deletion in WT strains. Gene deletions causing the most severe growth defect in combination with T4V in both screens are indicated with locus name, gene name and brief description.

Supplementary Figure S5.



Htz1 is eviction from the activated *GAL1* promoter is unaffected in the CONS strain.

(A) Western blot analysis of WT, CONS, and T4V strains expressing 6x HA tagged Htz1. Levels of Htz1-6HA and RNAP II (Rpb1) in the same samples are shown.

(B) ChIP analysis of Htz1-6HA occupancy at the *GAL1* promoter before and after induction with galactose in WT or CONS strains.

Supplementary Table S1. RNA-seq statistics for WT and T4V analysis.

Sample	No. of raw reads	No. of uniquely mapped reads
T4V replicate 1	4,216,055	3,601,290
T4V replicate 2	5,374,604	4,580,668
WT replicate 1	5,048,318	4,333,368
WT replicate 2	4,502,798	3,897,106

Supplementary Table S2. RNA-seq statistics for WT and CONS analysis.

Sample	**No. of raw reads**	**No. of uniquely mapped reads**
SC medium, CONS ("26p"), replicate 1	**8,974,658**	**2,804,521**
SC medium, CONS ("26p"), replicate 2	**10,825,427**	**3,353,291**
SC medium, WT, replicate 1	**6,436,125**	**1,813,413**
SC medium, WT, replicate 2	**6,778,954**	**1,935,444**
XY (rich) medium, CONS ("26p"), replicate 1	**9,905,275**	**3,015,817**
XY (rich) medium, CONS ("26p"), replicate 2	**4,563,665**	**1,416,412**
XY (rich) medium, WT, replicate 1	**10,244,166**	**3,249,571**
XY (rich) medium, WT, replicate 2	**8,221,832**	**2,357,730**

Supplementary Table S3. Yeast Strains Used in this Study All derivative strains were constructed for this study.

Strain	Genotype
W303a	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* **The following strains were derived from W303a.**
ERYM745A	*RPB1::RPB1-NatR*
ERYM746A	*rpb1::rpb1-CTD(CONS)$_{26}$-NatR*
ERYM747B	*rpb1::rpb1-CTD(T4V)$_{24}$-NatR*
ERYM748A	*htz1Δ ::kanMX*
ERYM779A	*ino80Δ900::kanMX HTZ1-6HA::Kl TRP1*
ERYM749A	*RPB1::RPB1-NatR HTZ1-6HA::Kl TRP1*
ERYM750A	*rpb1::rpb1-CTD(CONS)$_{26}$-NatR HTZ1-6HA::Kl TRP1*
ERYM751B	*rpb1::rpb1-CTD(T4V)$_{24}$-NatR HTZ1-6HA::Kl TRP1*
NYYM152A	*ybr184wΔ::kanMX*
NYYM153A	*ybr184wΔ::kanMX rpb1::rpb1-CTD(T4V)$_{24}$-NatR*
NYYM145A	*swr1Δ::kanMX*
NYYM146A	*swr1Δ::kanMX rpb1::rpb1-CTD(T4V)$_{24}$-NatR*
NYYM147A	*swc5Δ::kanMX*
NYYM148A	*swc5Δ::kanMX rpb1::rpb1-CTD(T4V)$_{24}$-NatR*
The following strains were derived from the BY4172 background.	
Y7092	*can1Δ::STE2pr-SpHis5 his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 lyp1Δ LYS2+*
NYYM115A	*rpb1::rpb1-CTD(CONS)$_{26}$-NatR* (derived from Y7092)
NYYM116C	*rpb1::rpb1-CTD(T4V)$_{24}$-NatR* (derived from Y7092)

MPK1/SLT2 links multiple stress responses with gene expression in budding yeast by phosphorylating Tyr1 of the RNAP II CTD

Nathan Yurko[1], Xiaochuan Liu[2], Takashi Yamazaki[1], Mainul Hoque[2], Bin Tian[2] and James L. Manley[1]

[1]Department of Biological Sciences, Columbia University, New York, NY 10027, USA

[2]Department of Microbiology, Biochemistry and Molecular Genetics, Rutgers New Jersey Medical School, Newark, New Jersey, USA

Summary

The RNA polymerase II (RNAP II) largest subunit C-terminal domain (CTD) consists of repeated YSPTSPS heptapeptides. The role of tyrosine-1 (Tyr1) remains incompletely understood, as, for example, mutating all Tyr1 residues to Phe (Y1F) is lethal in vertebrates but a related mutant has only a mild phenotype in *S. pombe*. Here we show that Y1F substitution in budding yeast resulted in a strong slow-growth phenotype. The Y1F strain was also hypersensitive to several different cellular stresses that involve MAP kinase signaling. These phenotypes were all linked to transcriptional changes by RNA sequencing, and we also identified genetic and biochemical links between Tyr1 and both transcription initiation and termination factors. Further studies uncovered defects related to the MAP kinase I (MPK1/SLT2) pathway, and we provide evidence that Slt2 phosphorylates Tyr1 in vitro and in vivo. Our study has thus identified Slt2 as a Tyr1 kinase, and in doing so provided links between stress response activation and Tyr1 phosphorylation.

Introduction

In eukaryotes, the multisubunit enyzme RNA polymerase II (RNAP II) is responsible for transcription of all mRNAs as well as numerous noncoding RNAs. The largest subunit of RNAP II, Rpb1, contains a C-terminal domain (CTD) consisting of 26 (in yeast) to 52 (in vertebrates) heptad repeats, the consensus sequence being Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 (Allison et al. 1985; Corden et al. 1985; Corden 1990; Liu et al. 2010; reviewed by Hsin and Manley 2012). The non-proline residues are all subject to phosphorylation and de-phosphorylation during the transcription cycle, and the different residues, and their modification, appear to serve distinct roles related to the stage of transcription (reviewed in Phatnani and Greenleaf 2006; Eick and Geyer 2013; Napolitano et al. 2014; Conaway and Conaway 2015). The CTD thus functions in essentially all aspects of the synthesis and processing of RNAP II transcripts, which includes not only initiation, elongation and termination of transcription, but also all the processing reactions, such as 5' capping, splicing and 3' end formation.

Of the phosphorylatable CTD residues, serines 2 and 5 have received the most attention. Ser5 phosphorylation is concentrated near the 5' end of genes, primarily as a result of the TFIIH-associated kinase Kin28 (Feaver et al. 1994), where it functions in recruitment of capping enzyme and a number of factors important for transcriptional elongation (Cho et al. 1997; Ng et al. 2003; Govind et al. 2007). Additionally, the multisubunit Mediator takes advantage of CTD flexibility and aids in Kin28 phosphorylating all CTD heptads by guiding it through the Mediator-PIC (pre-initiation complex) head module (Robinson et al. 2016). Ser5P is also directly recognized by Nrd1 of the NNS complex, which also consists of Nab3 and Sen1, to terminate

transcription of short RNAP II products, including cryptic unstable transcripts (CUTs), stable uncharacterized transcripts (SUTs) and small nucleolar RNAs (Arigo et al. 2006, Vasiljeva et al. 2008). The primary driver of Ser2 phosphorylation in yeast, Ctk1, is responsible for the bulk of Ser2P, which accumulates towards the 3’ end of genes (Mayer et al. 2010). Ser2P levels peak near 3' cleavage-polyadenylation sites, where termination factors such as Pcf11 and Rtt103 interact directly with the CTD, facilitating 3' processing and termination (Kim et al. 2004; Luo et al. 2006).

Determining the function of Ser2 and Ser5, as well as other residues in the heptad repeat, has benefited from the strategy of directly modifying the CTD. Besides direct substitution mutations (e.g., Ser-to-Ala), more recent efforts (Schuller et al. 2016; Suh et al. 2016) have altered the CTD to allow detection of phosphorylation patterns using mass spectrometry. These studies confirmed that all non-proline residues are phosphorylated, although replacing many individual residues (mostly Ser7) with lysine and arginine (enabling trypsin cleavage) produced a slow growth phenotype under cold stress (Suh et al. 2016). Extending the importance of Ser2 and Ser5, these studies found that Tyr1 and Thr4 phosphorylation was much less abundant compared to Ser2/5, although whether or not this was due to an effect of the other CTD modifications or to the relative importance of Tyr1/Thr4 under the growth conditions used is not clear.

The other phosphorylatable CTD residues (Ser7, Thr4 and Tyr1) appear to perform more specialized roles (reviewed in Yurko and Manley 2017). After Ser2 and Ser5, Ser7 phosphorylation has been well characterized, and is present on RNAP II on active genes (Chapman et al. 2007). Ser7 is also important for proper small nuclear RNA gene expression, such as U1/U2 snRNA 3' end formation, in mammals (Egloff et

al. 2007), and enables better recruitment of the functional human homolog of Ctk1, P-TEFb, to the CTD (Mayer et al. 2010; Czudnochowski et al. 2012). Thr4 has been found to have differing roles depending on the organism. In chicken cells expressing a T4V derivative, histone mRNA 3' processing was found to be defective (Hsin et al. 2011), and both the chicken T4V and human T4A (Hintermair et al. 2012) were lethal. In budding yeast, Thr4 is required for the induction of certain classes of genes, and is necessary for proper variant histone H2A.Z eviction and induction of PHO and GAL metabolic genes (Rosonina et al. 2014). Recent work also revealed a role in budding yeast for Thr4 in splicing and Rtt103 recruitment (Harlen et al. 2016).

The function of Tyr1 has only more recently been investigated. Chicken DT40 cells expressing a Y1F derivative were inviable, and also revealed that a function of Tyr1, and its phosphorylation, was stabilization of the CTD (Hsin et al. 2014). Tyr1 was also required for efficient turnover of certain unstable transcripts, such as upstream antisense RNAs (uaRNAs), and Tyr1 phosphorylation was indeed elevated on regions encoding uaRNAs. Very similar results were observed in human cells (Descostes et al. 2014). A Y1F derivative was also found to be lethal in budding yeast (West and Corden 1995), and Tyr1P patterns along yeast genes are similar to Ser2P (Mayer et al. 2012). Tyr1 phosphorylation is also important for histone chaperone Spt6 binding (Burugula et al. 2014), and the presence of Tyr1P impairs Nrd1 and Rtt103 interaction with CTD peptides, suggesting a role in modulating termination (Lunde et al. 2010; Mayer et al. 2012). Surprisingly, in fission yeast Tyr1 can be mutated with no loss of viability, although cold- and iron-sensitivity phenotypes were noted (Schwer and Shuman 2011; Schwer et al. 2012, 2014). While evidence exists that Tyr1 can be phosphorylated by

the c-Abl kinase in mammals (Baskaran et al. 1993), the responsible kinase(s) in yeast, where tyrosine phosphorylation is rare, is unknown.

Here, we investigate the role of Tyr1 and Tyr1P in *S. cerevisiae*. We first generate a strain that expresses a full-length Y1F form of the CTD and show that the strain is viable but very slow growing under normal conditions. We identify genetic links between the Y1F mutant and Mediator, and found that the Y1F phenotype can be partially suppressed by removing Srb10/11 subunits. We also found that several stress-response genes dealing with cell wall, oxidative, and related stress responses, as well as iron response (Schwer et al. 2014), are dysregulated in Y1F cells, and that the cells are hypersensitive to stressors that specifically target these pathways. Bringing these observations together, we demonstrate that the specific MAP kinase involved in these processes, Slt2, phosphorylates Tyr1 in vitro, and that in vivo Slt2 is necessary for full Tyr1 phosphorylation and modulates Tyr1P levels during stress responses. These effects are accompanied by defects in transcription termination factor recruitment, as Y1F disrupts Rtt103 binding to the CTD and impairs Nrd1 recruitment to chromatin. Our study thus links Tyr1 with Slt2-mediated stress responses, at least in part via anti-termination, and reveals Slt2 as a Tyr1 CTD kinase.

Results

Mutation of all CTD Tyr1 residues results in severe growth defects

In this study, we wished to gain more insight into the function of CTD Tyr1 in budding yeast. In our previous work, we found that replacing the CTD of chicken Rpb1 with a Tyr1-to-Phe (Y1F) derivative resulted in lethality (Hsin et al. 2014), consistent

with an earlier study in budding yeast (West and Corden 1995), but in contrast to similar experiments with *S. pombe* (Schwer et al. 2014). All of these studies, however, employed truncated CTDs, and we therefore constructed a yeast Y1F strain containing 26 FSPTSPS repeats (Figure 1A, see Experimental Procedures). As a control, an all-consensus YSPTSPS CTD was inserted into the same background, as we had done in our previous analysis of Thr4 (Rosonina et al. 2014).

We next determined several properties of the Y1F strain, which we will refer to simply as Y1F. Comparing the growth of the parental strain (WT), consensus (CON), and two separate Y1F (Y1F 2A and 2B) strains revealed that Y1F displayed a severe slow-growth defect on both rich and synthetic complete media (Figure 1A). This phenotype was somewhat alleviated by growing at 37^{o}C, but, consistent with related results in fission yeast (Schwer et al. 2014), Y1F was cold sensitive and unable to grow at lower temperatures (Figure 1A). Tetrad analysis showed that the slow growth phenotype was specific to the Y1F mutation, suggesting the absence of suppressors that would prevent lethality (Figure S1A), and the slow growth defect was also observed in liquid media (Figure S1B). Cell-cycle analysis showed that Y1F cells are defective in exiting G1/G0 and entering G2 (Figure 1B).

We next examined accumulation of Rpb1 and its phosphorylation status by Western blot (Figure 1C). As expected, the WT strain showed evidence of Tyr1P as well Ser2P and Ser5P. However, Y1F Rpb1, while as expected not detected by an anti-Tyr1P antibody (3D12), was also not recognized by antibodies specific for unphosphorylated consensus repeats (8WG16) or Ser2P (3E10). This likely indicates an inability of these antibodies to recognize the Y1F epitope (see below). Indeed, a

different Ser2P antibody (H5) produced a much stronger signal with the Y1F CTD, although whether this represents a bona fide increase in Ser2P, an increase in Ser2P with Ser5P (see Phatnani and Greenleaf 2006; Chapman et al. 2007; Devaiah et al. 2012), or differential recognition due to an altered epitope, is not known. Ser5P levels appeared unaffected using the antibody 3E8. Finally, analysis with an antibody (y-80) that recognizes an epitope in the first 80 amino acids of Rpb1 showed that Y1F Rpb1 accumulated to levels comparable to WT and CON Rpb1, and was full length. This contrasts with the situation in vertebrate cells, where Y1F mutations resulted in CTD destabilization (Descostes et al. 2014; Hsin et al. 2014). Additionally, levels of the hyperphosphorylated Rpb1 IIo isoform were also similar, suggesting that the apparent differences in Ser2P most likely reflected differences in epitope recognition by the phospho-specific Abs. Supporting the conclusion that Ser2 phosphorylation was unaffected, a co-immunoprecipitation (co-IP) assay with WT or Y1F Rpb1 and Ctk1 (the primary Ser2 kinase) revealed no changes in Ctk1 association with Rpb1, providing evidence that Ser2 phosphorylation by Ctk1 was not affected by the Y1F mutation (Figure 1D).

Tyr1 interacts genetically with Mediator subunits SRB9/10/11

To investigate what pathways are affected by the Y1F mutation, we carried out a synthetic genetic array assay (SGA) (Tong et al. 2001, Rosonina et al. 2014). Due to the slow growth phenotype of the Y1F mutant, we performed the SGA as a suppressor assay, using a 5FOA-URA3 system with a plasmid-borne WT copy of Rpb1, comparing Y1F and CON strains (Poschke et al. 2012). Genes related to transcription found in each of two independent suppressor SGA analyses are listed in Figure 2A (the

complete list is in Dataset S1), and those that best suppressed the Y1F slow growth phenotype are shown in Figure 2B. Sample tetrad dissections are shown in Figure S2A, and suppression of the Y1F growth defect could also be observed in liquid cultures, (Figure S2B). Interestingly, subunits of the Srb10/11 Mediator kinase module, including its scaffolding protein (Srb9), suppressed the Y1F phenotype when deleted, and were responsible for three of the four strongest interactions detected. The Srb10/11 kinase/cyclin pair is known to inhibit transcription of numerous genes (many as part of the general stress response) and can phosphorylate transcription factors such as Gcn4 to dampen activation, in tandem with other post-translational modifications such as ubiquitination and SUMOylation (Chi et al. 2001; Nelson et al. 2003; Bose et al. 2005; Rosonina et al. 2012; Allen and Taatjes 2015). The other strongest hit, Ubp8, is a component of the SAGA complex, a histone acetyltransferase that functions as a transcriptional activator. Ubp8 is a ubiquitin-specific protease required for deubiquitination of histone H2B (Lee et al. 2005; Koutelou et al. 2010; Wilson et al 2011).

We next investigated the significance of the Srb10/11 SGA suppressor result. For this, ChIP assays were conducted using 3HA-tagged strains in both WT and Y1F backgrounds. Four genes regulated by Srb10/11 (*FKS2*, encoding a cell wall enzyme [de Virgilio 2012], cell-cycle coordinator *CLN2* and meiosis regulator *IME1* [Ohkuni and Yamashita 2000], and *PMA1* [Cai et al. 2006]) and two genes not subject to Srb10/11 regulation *(ZIM17* and *PYK1*) were analyzed. We detected a roughly two-fold increase in promoter occupancy of Srb10/11 on all four target genes in the Y1F strain, while occupancy on the genes without a link to Srb10/11 was not affected (Figure 2C). This

suggests genes regulated by Srb10/11, which include stress response and cell-cycle genes (Dynlacht 1997; Krasley et al. 2006), are especially sensitive to Y1F, possibly explaining the Y1F cell-cycle defect. We note that although *SRB* mutations have long been known to suppress truncated CTD derivatives (Hengartner et al. 1995; Yuryev and Corden 1996), suppression of a full-length Y1F CTD likely reflects a distinct mechanism reflecting enhanced Srb10/11 promoter occupancy. These findings are discussed in more detail below.

Tyr1 is required for proper expression of stress-mediated genes

We next wished to extend our analysis by examining directly how Y1F alters global gene expression. To this end, we performed mRNA sequencing (RNA-seq) using 3'READS (Hoque et al. 2013) with WT and Y1F strains grown in rich medium. This method allowed us to determine possible effects of the Y1F on alternative polyadenylation as well as mRNA levels. In addition to a general shift favoring distal poly(A) site usage (Figure S3), changes in expression of multiple genes were observed. Expression of 5,925 genes are tabulated in Dataset S2. Transcript levels of 306 genes were significantly increased, while 329 were reduced (adjusted P-values less than 0.05). GO analysis was then performed, with the top associated terms listed in Figure 3A. Consistent with the SGA and/or previous work (Schwer et al. 2014), terms enriched from this analysis include cell wall stress, iron homeostasis, and processes related to oxidoreductive stress. Correlations between samples, which was as high as 0.94 for significant genes, are shown in Figure S4. Additionally, a significant upregulation of both CUTs (P-value of 7.49 x 10-12) and SUTs (P-value of 2.23 x 10-5) was detected (Figure S5). RT-PCR analysis of a sample (12) of the genes with altered expression validated

the RNA-seq results (Figure 3B). Among these were FKS2 (associated with multiple stress responses), SIT1 (iron stress response), and CAP2 (DNA damage response).

We also compared our data with an existing Tyr1P ChIP dataset (Mayer et al. 2012). Out of our 630 genes displaying altered expression in Y1F cells, 498 were also found in the Mayer dataset (222 up-regulated, 276 down-regulated). Strikingly, the up-regulated genes had a Tyr1P ChIP signal weaker than average (P-value of 4.45 x 10-8) whereas the down-regulated genes had a Tyr1P signal stronger than average (P-value of 6.60 x 10-10) (summarized in Figure S6). These results provide a strong correlation between genes with altered expression in Y1F cells and levels of Tyr1 phosphorylated RNAP II.

The GO analysis described above suggested that expression of genes involved in several specific stress responses was altered in Y1F cells. Notably, however, the expression changes detected were not due to a generalized stress response (Gasch et al. 2000; O'Duibhir et al. 2014), as there was poor overlap between our Y1F data and the Extended Stress Response (ESR) gene set identified by Gasch et al. (11 out of 138 genes from the ESR overlapped). Interestingly, though, we noted that cell wall and iron stress responses both require MAP kinase signaling, specifically the MAPK Slt2, for proper function, and the DNA damage/oxidative stress response does as well (Levin 2005, 2011; Pujol-Carrion et al. 2013; Soriano-Carot et al. 2012). We therefore next examined possible links between Y1F and stress responses involving MAP kinases. To this end, we compared growth of Y1F and WT in medium containing compounds that perturb MAP kinase-associated pathways (Figure 3C). These include calcofluor white (CFW), phleomycin, hydroxyurea, FK506, and rapamycin. Strikingly, with three of the

compounds (CFW, phleomycin, and hydroxyurea), Y1F was completely inviable while WT growth was unaffected, and both rapamycin and FK506 slowed growth of Y1F significantly more than WT. Additional stress pathways not known to involve Slt2 (e.g., galactose metabolism, high salt and low phosphate conditions) were also tested and found not to significantly alter growth of Y1F compared to WT (Figure S7). These observations, together with the RNA analysis above, indicate that although Tyr1 is not required for normal expression of most genes under normal growth conditions, a subset of MAP kinase-induced genes, and those important for MAP kinase function (such as MKK1), require Tyr1. This in turn implicates Tyr1, and perhaps its phosphorylation, in the response to MAPK-related stresses.

Slt2 phosphorylates Tyr1 in vitro and in vivo

The above data suggesting that MAPK signaling pathways are defective in Y1F cells raises the possibility that a kinase in this pathway may naturally function to phosphorylate Tyr1. Two candidates are Slt2 and Hog1, both of which are Ser/Thr kinases that have also been shown capable of Tyr phosphorylation (Levin-Salomon et al. 2009; Maayan et al. 2012). Slt2 is a component of the PKC1 signaling pathway (reviewed by Pearson et al. 2011), which regulates the cellular response to several stresses, such as heat shock, cell wall stress, and DNA damage, through MAP kinase cascades (Soriano-Carot et al. 2012). Slt2 acts in transcription through phosphorylation of transcription factors (TFs), such as the Swi4/Swi6 complex, also known as SBF (Koch et al. 1996; Kim and Levin 2010; Kahana-Edwin et al. 2013). Activated Slt2 is recruited to promoters to increase transcription by phosphorylating TFs as well as by mediating destruction of Srb11, thereby further increasing transcription (Madden et al.

1997; Baetz et al. 2001; Jung et al. 2002; Levin 2011). Furthermore, Slt2 promotes transcription of stress related genes not only by phosphorylation of SBF and other TFs, but also by blocking association of the NNS complex, preventing early transcription termination (Kim and Levin 2011; see also Kim et al. 2008). Hog1 functions in the cellular response to osmotic stress, and is homologous to p38 and JNK kinases in mammals (reviewed in Brewster and Gustin 2014). Hog1 controls expression of more than 300 genes by association with the RNAP II complex at promoters (Rep et al. 2000).

In light of the above, we tested the possibility that Slt2 and/or Hog1 have Tyr1 kinase activity. We first used an in vitro kinase assay (adapted from Campbell 2014) to determine if either kinase can phosphorylate a GST-CTD derivative. For this, we generated 3HA-tagged versions of Slt2 and Hog1 to allow us to immunoprecipitate (IP) these kinases from cell extracts. We IPed HA-tagged Slt2 and Hog1 from WT cells (after activating the kinases through exposure to the appropriate stress conditions, i.e. CFW or NaCl, respectively), incubated them with a purified GST-CTD fusion protein, and analyzed phosphorylation by Western blot with CTD phospho-specific antibodies (Figure 4A). Importantly, we found that Slt2, but not Hog1, can indeed phosphorylate Tyr1. Consistent with previous data (Akhtar et al. 2009; Chasman et al. 2014), Slt2 and Hog1 also phosphorylated Ser2 and Ser5. As controls, strains carrying kinase-dead (KD) HA-tagged derivatives of these two kinases (see Kim et al. 2008; Irqeba et al. 2014) were prepared and used in the in vitro kinase assay. Neither KD derivative showed any kinase activity (Figure 4A). As additional controls for specificity, we constructed and analyzed strains expressing 3HA-tagged Kin28 and Ctk1, which

phosphorylated Ser2 and/or Ser5 but not Tyr1 (Figure S8). Our data thus provide strong evidence that Slt2 possesses CTD Tyr1 kinase activity.

We next investigated whether Slt2 can affect Tyr1 phosphorylation status in vivo. For this, we first subjected WT cells to MAPK pathway-related stresses, specifically cell wall stress resulting from exposure to CFW and DNA damage stress from phleomycin exposure. Strikingly, Western blot analysis of cell lysates revealed a significant increase in Tyr1P following CFW exposure (3.0-fold after one hour, normalized to Rpb3 levels), but not of Ser2P or Ser5P (Figure 4B, quantified in Figure 4C). While phleomycin treatment resulted in a decrease in Tyr1P after 15 minutes, a net 1.5-fold increase was observed after one hour (Figure 4B). Slt2 levels following CFW and phleomycin exposure were increased relative to the DMSO alone control (following a decrease after 15 minutes in the case of phleomycin), while Hog1 levels were unchanged by either treatment (Figure 4B). To determine whether the increased Tyr1P was in fact due to Slt2, we generated strains carrying deletions in *SLT2* (SΔ), *HOG1* (HΔ), or both genes (SHΔ). We also analyzed the kinase-dead strains described above (SM and HM). WT cells were grown under normal conditions (*SLT2* mutant strains are inviable in the presence of CFW [van Voorst et al. 2006]), and lysates analyzed by Western blot (Figure 4D). Importantly, Tyr1P was significantly reduced in the SΔ, S-KD and SH1Δ strains (1.8- , 1.9- , and 3.0- fold, respectively), but not in the HΔ and H-KD strains. These findings strongly support our in vitro data that Slt2 is a Tyr1 kinase, although as Tyr1P was not completely abrogated in the *SLT2* mutant strains, another unknown kinase may also target Tyr1.

We extended this analysis to another MAPK-related stress, heat shock. Consistent with earlier studies (Kamada et al. 1995; Kim et al. 2010), and similar to what we observed following CFW exposure, heat shock resulted in an increase in overall Slt2 levels (Figure 4E). Importantly, heat shock induced a 3.4-fold increase in Tyr1P, while Rpb1 levels and Ser2P/Ser5P were unchanged. Due to lethality of heat exposure for *SLT2* mutants (Torres et al. 1991), we were unable to measure Tyr1P levels under heat stress in the deletion and kinase-dead strains described above.

We next examined if Slt2 associates with RNAP II, and if so how this might be affected by Y1F, and whether this correlates with Tyr1P levels. To this end, we performed coIP assays with extracts from WT and Y1F cells expressing 3HA-tagged Slt2 (Figure 5A). Using the y-80 antibody for IP, we found that Slt2 was coIPed with RNAP II from WT cell extracts, but not from Y1F extracts. (DNAse treatment did not alter the results; data not shown). To extend these results, we prepared extracts from WT cells expressing 3HA-tagged Slt2, either growing normally or treated with CFW to activate the kinase, and from normally growing cells expressing 3HA-tagged KD Slt2. Extracts were IPed with anti-HA antibodies and Western blots probed with anti-Tyr1P, -Rpb1, -Ser2P and -Rpb3 antibodies. Strikingly, Slt2-associated Tyr1P levels were 2.3-fold higher in the extracts from activated relative to normal cells, while 0.6-fold lower in the kinase-dead strain (Figure 5B); no changes in Rpb1, Ser2P, or Ser5P levels were detected. Taken together, our findings establish Slt2 as a CTD Tyr1 kinase, and in doing so link Slt2-mediated Tyr1 phosphorylation with the cellular stress response.

Y1F alters Nrd1 and Rtt103 interactions with RNAP II

We next wished to investigate how Tyr1, and its phosphorylation by Slt2, might affect target gene expression. Since Tyr1 phosphorylation and Slt2 activity are both known to influence association of certain termination factors with the transcription machinery (Satoh et al. 2009; Kim and Levin 2011; Heidemann and Eick 2012; Mayer et al. 2012; Porrua and Libri 2015), and the Y1F strain shows a general shift to distal poly(A) site usage and upregulation of both CUTs and SUTs (Figures S3 and S5), we examined if recruitment of such factors is altered by Y1F. To do this, we attempted to generate 3HA-tagged versions of Nrd1, Rtt103 and Pcf11, three proteins previously shown to interact with the CTD peptides in a Tyr1P-sensitive manner (Mayer et al. 2012), in WT and Y1F backgrounds. While we were unable to tag Pcf11 in the Y1F strain, as Y1F and the Pcf11 tag were synthetically lethal, we did generate Nrd1- and Rtt103- tagged strains. These were then used in coIP experiments, using the y-80 antibody. Consistent with earlier predictions (Mayer et al. 2012), Y1F Rpb1 showed a 7.3-fold increase in Nrd1 binding (Figures 6A). In contrast, Rtt103 levels in the co-IP were significantly reduced in the Y1F strain (Figure 6B). Although not entirely as expected (Mayer et al. 2012), a possibility consistent with previous studies (Lunde et al 2010; Mayer et al. 2012) is that Rtt103 requires the hydroxyl group of Tyr1, which phosphorylation as well as Phe substitution disrupts, to bind the CTD efficiently. Ser2 (Lunde et al. 2010) and Thr4 (Jasnovidova et al. 2017) are also required for this association.

We next examined Nrd1 occupancy in WT and Y1F strains on genes known to be Nrd1/NNS targets. Using ChIP, we found unexpectedly that Nrd1 levels, normalized to Rpb1 occupancy, decreased in the Y1F strain near promoters for *FKS2*, *IMD2* and

HRP1 (all previously shown to be Nrd1/NNS targets; Kim and Levin 2011, Steinmetz et al. 2006), as well as for *ZIM17* (one of the most significantly downregulated genes in Y1F; see Figure 3A). In contrast, Nrd1 occupancy on *PMA1*, not known to be an NNS target, was not affected (Figure 6C). This decreased Nrd1 occupancy on target promoters is consistent with the general upregulation of CUTs and SUTs we observed in Y1F cells (see above).

To gain more insight into how Y1F effects Nrd1 occupancy on target promoters, we examined Nrd1 and Slt2 occupancy on a target gene following heat shock. To this end, we generated paired Nrd1 and Rpb1 ChIP samples from control and heat-shocked WT and Y1F cells, and analyzed both Nrd1 and Slt2 occupancy on *FKS2*, which has served as an important model gene for both NNS function and MAPK signaling (Kim et al. 2008; Kim and Levin 2011). We found that Nrd1 occupancy relative to Rpb1 at the promoter region was reduced following heat shock in WT cells, as expected (Kim and Levin 2011). However, upon heat shock Nrd1 occupancy relative to Rpb1 in Y1F cells increased significantly (2.5-fold; Figure 6D), indicating that Tyr1, and likely its phosphorylation, is required for loss of NNS following heat shock. Supporting a role for Slt2 in this process, while heat shock resulted in Slt2 recruitment to promoters in WT cells, as observed previously (Kim and Levin 2011), this was not observed in Y1F cells (Figure 6E), consistent with the co-IP results shown above. Together, and as discussed in more detail below, our results implicate Tyr1 and its phosphorylation by Slt2 in controlling Nrd1/NNS function during the stress response.

Discussion

Our studies have provided new insights into the function of RNAP II CTD Tyr1 residues in budding yeast. While previous analyses of Tyr1 function in vertebrate cells (Descostes et al. 2014; Hsin et al. 2014) point to significant differences between yeast and higher eukaryotes, one notable similarity is that this highly conserved residue plays relatively specific roles in gene expression. We have shown here that Tyr1 is required for expression of a subset of stress-inducible genes, notably MAP kinase-associated genes. Importantly, we found that Tyr1 is phosphorylated by Slt2 in vitro, and that the levels of Tyr1P in cells can be modulated by exposure to specific stresses known to require Slt2 function, or by inactivating the kinase. Given previously defined roles of both Tyr1 and Slt2 in anti-termination, we provided evidence for Slt2 and Nrd1 co-regulation of stress-related target transcripts, showing that Tyr1 provides a nexus at which stress response and anti-termination meet to influence transcription. Below, we discuss how Slt2 activity and Tyr1 serve as the connection for these processes in budding yeast, as well as the varied roles Tyr1 and its phosphorylation play throughout evolution.

Our data implicate Tyr1 in the proper function of the cell wall integrity pathway, which provided the first suggestion that Slt2 may function as a Tyr1 kinase. Upon sensing cell wall stress, such as through activation of the cell wall sensor Mid2, the Rho1 effector pathway is activated, which in turn activates the MAP kinase Pkc1, triggering a kinase cascade resulting in activation of Slt2 (Posas et al. 1998; Levin 2005). Slt2 has a diverse number of nuclear targets, many of them transcription factors such as Swi4 of the SBF complex (Baetz et al. 2001). Activated Slt2 is characterized by a phosphorylated TxY activation loop, which changes the conformation of the protein to

allow for sustained phosphorylation of targets (Chen and Thorner 2007, Truman et al. 2009). While previous investigations of Slt2 focused on serine phosphorylation, removal of the C-terminus of Slt2 enables autophosphorylation of the TxY activation loop itself, a method of regulation that has in fact been observed in higher eukaryotes (Levin-Salomon et al. 2009; Goshen-Lago et al. 2016; Smorodinsky-Atias et al. 2016).

The involvement of Srb10/Srb11, two of the suppressors identified in the SGA, in stress responses has been known for some time. The kinase/cyclin pair is a target of Slt2, with Slt2 phosphorylation leading to degradation of Srb11 (Jin et al. 2014; Strich and Cooper 2014). Previous genetic studies have linked Srb10/11 function to Slt2 response, as this module of Mediator is linked to heat, oxidative and anaerobic stress through regulation of the TF Skn7 (Arias et al. 2011). The association between oxidative and cell wall stresses and Srb11 has been known longer, since it was found that Ask10 (“Activator of Skn7”) directly mediates the destruction of Srb11 in response to oxidative stress using the Pkc1 pathway (Cohen et al. 2003). Additionally, increased expression of certain stress response genes, particularly iron response, upon phosphorylation of Mediator subunit Med2 by Srb10, has been observed (van de Peppel et al. 2005), which is consistent with our ChIP and RNA analyses. Our data that Tyr1 is required for loss of Srb10/11 from target genes extends the role of Slt2 in the cell wall integrity (CWI) pathway, which previously was not known to target RNAP II directly. Previous work established Slt2's role in facilitating Srb11 translocation from nucleus to cytoplasm (Jin et al. 2014) subsequent to its degradation by the proteasome (Cooper et al. 1997, Krasley et al. 2006, Strich and Cooper 2014). This function is impaired in Y1F cells, as Srb10/11 occupancy on several target promoters increased in these cells, and Slt2 was

not recruited to at least one of these genes (*FKS2*), likely reflecting its defective interaction with Y1F RNAP II.

The involvement of Slt2 in the expression of stress response-related genes has been well documented. Our data extend these findings by showing that Tyr1 phosphorylation by Slt2 is critical in several stress responses. Besides the canonical cell wall integrity pathway, as exemplified by sensitivity to calcofluor white (reviewed in Levin 2005, 2011), Slt2 is also involved in the response to DNA damaging agents such as hydroxyurea and phleomycin (Soriano-Carot et al. 2012), and functions along with calcineurin upon exposure to FK506 (Mizunuma et al. 1998). Additionally, in concert with TORC1 signaling, Slt2 helps stabilize the cell's response to rapamycin (Moreno-Torres et al. 2015), and also serves as a target of TORC1 phosphorylation upon caffeine exposure (Kuranda et al. 2006). Importantly, many of the genes with altered expression in Y1F were already known to be regulated by Slt2, as part of the above stress responses. In a related Y1F *S. pombe* strain, a similar response to phleomycin was observed, which was the result of iron uptake pathway upregulation generating excess intracellular iron, potentiating phleomycin activity (Schwer et al. 2014). Our study extends this finding significantly, not only by expanding the role of Tyr1 in the stress response to most Slt2-dependent responses, but also by involving Tyr1-dependent interactions directly in Slt2 function. We thus expect additional Slt2-dependent stress responses to involve Tyr1, either through direct Tyr1 phosphorylation or indirectly through regulation of other Slt2 targets.

Previous studies of Tyr1 revealed differences in the distribution of Tyr1P between species, as well as the importance of the residue to termination factor recruitment.

Genome-wide ChIP analyses in *S. cerevisiae* showed that Tyr1P reaches a peak at the 3' end of genes, although it begins to accumulate towards the 5' end, much like Ser2P (Mayer et al. 2012). This is in contrast to human cells, which show a peak at the 5' end of genes as well as in the antisense direction, and at transcriptional enhancers (Descostes et al. 2014). Further differentiating yeast and human cells, there is no clear human Nrd1 homolog, and the processing of snRNAs is done differently through the Integrator complex (Baillat et al. 2005, O'Reilly et al. 2014). It would make sense, then, that the functions associated with Tyr1 differ between yeast and human cells; the proposed function of Tyr1P in budding yeast (preventing termination factors Nrd1, Pcf11 and Rtt103 from prematurely binding to the CTD) seems to be yeast specific (Mayer et al. 2012). The involvement of Tyr1 in the stress responses we have shown here, mediated in part by the Nrd1/NNS complex, may also be yeast specific, but further studies will be needed to investigate this.

Our data provide new insight into the CTD interaction with the NNS complex. The substantial increase in Nrd1 binding to the Y1F CTD we observed is consistent with previous studies showing that Tyr1P-containing CTD peptides bound Nrd1 less strongly than unmodified peptides (Mayer et al. 2012). The finding that upon cell stress, Slt2 is induced and recruited to coding regions as well as promoters of specific genes links Slt2 with NNS function, as NNS association with elongating RNAP II drops with Slt2 induction (Kim and Levin 2011). While the presence of Slt2 on target genes was suggested to be sufficient for anti-termination (Kim and Levin 2011), our data provide evidence that Slt2 phosphorylation of Tyr1 is critical for this process by facilitating dissociation of NNS. Importantly, our findings that Tyr1P levels increase in response to

stress link Tyr1 phosphorylation by Slt2 with activation of stress-related genes. Based on these findings, we propose a model (Figure 7) that incorporates not only Slt2, Tyr1P and NNS, but also Srb10/11 and SBF, known targets of Slt2 (Jin et al. 2014, Kim et al. 2008) now linked to CTD Tyr1, into activation of stress response genes. While some of our observations were obtained from single-gene experiments, and thus may not apply in all circumstances, we believe the robust responses from our model system are indicative of more general mechanisms.

The significance of tyrosine phosphorylation in budding yeast is not entirely clear. *S. cerevisiae* encodes no typical tyrosine kinases, and the relative levels of TyrP are extremely low (<0.1% total phosphorylation; Modesti et al. 2001; Chi et al. 2007). Previous reports of Tyr phosphorylation have typically involved specialized functions, such as phosphorylation of heat shock protein Hsp90 by Swe1 (Mollapour et al. 2010). Tyr phosphorylation also occurs in the CWI pathway, as Mkk1 phosphorylates the Thr and Tyr residues of the Slt2 activation loop (Martín et al. 2000). Other dual-specificity kinases have been identified, including Hrr25 (Hoekstra et al. 1994), Yak1 (Kassis et al. 2000) and Kns1 (Lee et al. 1996). Our study adds Slt2 to this list, and establishes an important role for it in transcriptional control. Since inactivation of Slt2 did not abolish Tyr1 phosphorylation in vivo, it is likely that Tyr1 can be targeted by additional kinase(s), and it will be of interest to determine whether such kinases come from this group of dual-specificity kinases.

It remains to be seen how our findings on Tyr1 function extend to other organisms. As mentioned above, the corresponding Y1F mutation in *S. pombe*, despite the truncated CTD, did not cause nearly as strong a growth defect as the *S. cerevisiae*

derivative we analyzed (Schwer and Shuman 2011; Schwer et al. 2014). Notably, the *S. pombe* Y1F derivative retained a Tyr1-containing 4-repeat “rump,” offering at least a partial explanation for the relatively mild growth defects observed. The *S. pombe* and *S. cerevisiae* derivatives did however show similarities, including sensitivity to cold and phleomycin, and it is possible that future studies will reveal defects in additional stress pathways. It will also be of interest to determine whether the *S. pombe* Slt2 homologue, Pmk1, which functions in pathways similar to Slt2 (Toda et al. 1996, Madrid et al. 2006), is a Tyr1 kinase. Regulation of transcription by MAP kinase pathways also extends to higher eukaryotes, as homologs of Slt2, such as Erk1/2, were found to be recruited to chromatin in a manner similar to Slt2, on similar stress-induced genes (Pokholok et al. 2006; Yang et al. 2013; Mikula et al. 2016). Indeed, Erk1/2 are well known to have CTD Ser5 kinase activity (Trigon et al. 1998; Bonnet et al. 1999), although there is no evidence that they can target Tyr1. While there are significant differences between yeast and human cells in their response to stress (Verghese et al. 2012), it will be of interest to determine whether additional mechanisms evolved in human cells to amplify or refine signaling through Tyr1. In any event, our results show that Tyr1 and its phosphorylation by Slt2 play an important role in regulating transcription in response to stress in budding yeast, extending the complexity of CTD function in control of gene expression.

Acknowledgments

We thank Elizabeth Miller, Silvere Pagant, Rodney Rothstein and Robert Reid for sharing yeast strains, and members of the labs of JLM, BT, Elizabeth Miller and

Songtao Jia for technical help. We also thank David Levin for sharing the SLT2-K54R construct, as well as Yuqi Wang for sharing the HOG1-K52R strain. This work was supported by NIH grant R35 GM118136 to JLM, and NMY was partially supported by NIH training grant 5T32GM008798.

Experimental Procedures

Yeast strains and growth media

Yeast strains used are listed in Table S1. Y1F and CONS strains were constructed by PCR amplification of corresponding CTD constructs generated for use in previous studies (Hsin et al. 2011, Hsin et al. 2014), followed by transformation into haploid (for CONS) or diploid (for Y1F) S288C or BY4741/2 strains. The diploid strains carrying the Y1F CTD along with NAT-MX cassette (see Tong and Boone 2006, Rosonina et al. 2014) 3' of the *RPO21* gene were then confirmed by PCR, sporulation and tetrad dissection. The CONS strain was also marked with the NAT-MX cassette 3' of the RPO21 gene. Media containing FK506 (75 to 150 ug/uL), caffeine (5 mM), rapamycin (10 nM), phleomycin (1 ug/uL), or calcofluor white (3.5 ug/uL, kept in light-block environment) were prepared by adding the indicated amounts to rich media; hydroxyurea (50 mM) was added to synthetic complete media. SGA screen, yeast spot assays and media used were as previously described (Rosonina et al. 2009, 2010). Cell cycle analysis was performed according to Zhang and Siede (2004).

ChIP, RNA and protein analyses

For ChIP analyses, 50-mL cultures were grown and cells were lysed by bead-beating three times for one min each with one min in between in a cold room, as

previously described (Rosonina et al. 2014). After centrifugation, extracts were IPed with the appropriate antibody overnight, washed and treated with pronase, and DNA was extracted using phenol/chloroform. Samples were processed and analyzed, and data quantified, as previously described (Rosonina et al. 2009). Radioactive semi-quantitative PCR was performed as previously described (Rosonina et al. 2012), using a region of chromosome VII as internal control. Statistical analyses were performed using a two-tailed Student's t-test, with p-values indicated above bars in the graph. Error bars in graphs represent standard deviations of at least three experiments.

Western blot and co-IP assays were as previously described (Rosonina et al. 2009, 2010). Briefly, cell pellets from liquid cultures harvested between 0.5 and 1.0 OD_{600} were washed with IP buffer (see Rosonina 2009) and cells lysed using glass beads. Lysates were removed from glass beads then centrifuged at high speed (15,000 RPM, Eppendorf Centrifuge 5424) to pellet chromatin and debris. Supernatants were then centrifuged at high speed again and either diluted with 2X SDS loading buffer or IPed with sepharose beads and the appropriate antibody overnight at 4^{o}C before addition of sample buffer. DNAse treatment of samples was based on NEB protocol M0303, where 1 unit of DNAse I per 50 uL of protein extract was added, and samples were incubated at 37^{o}C for 10 minutes. This was performed before diluting samples with 2X SDS loading buffer. The reaction was stopped by adding 2X SDS buffer and boiling samples for 5 minutes.

Antibodies used for ChIP and western blotting were HA (ABM; ChIP and CoIP), HA (12AC5, a gift from Elizabeth Miller; Western), GAPDH (Sigma), MPK1 (sc-374434) and HOG1 (sc-165978) (Santa Cruz Biotechnology), Rpb3 (Biolegend), Y1P 3D12

(Active Motif), Ser2P (3E10, a gift from Dirk Eick), Ser5P (3E8; Millipore) and Rpb1 (y-80; Santa Cruz). Primer sequences used for cloning, ChIP and RT-PCR analyses are available upon request.

Random hexamers were used with total RNA for RT-PCR analyses, performed in triplicate.

RNA sequencing and data analysis

RNA sequencing was performed using Y1F and WT samples, in duplicate, and 3'READS data were analyzed as previously described (Hoque et al. 2013). Briefly, 3'READS reads were mapped to the *S. cerevisiae* genome (sacCer3) using bowtie2 with local mode. Uniquely mapped reads (with MAPQ score > 10) that had at least two additional 5' Ts after genomic alignments were assigned to genes by gene models from UCSC database (Tyner et al. 2017). CUTs and SUTs were annotated as described (Xu et al. 2009). DESeq was used to identify differentially expressed genes (Love et al. 2014). Dysregulated genes were selected by FDR ≤ 0.05.

Gene Ontology analysis

Gene Ontology (GO) annotations were obtained from the Gene Ontology Consortium (Gene Ontology Consortium, 2015). GO entries were tested for significance of association with regulated genes using the hypergeometric test (Grossmann et al. 2007). GO terms associated with more than 2,000 genes were discarded as too generic. To reduce redundancy, each represented GO term was required to have at least 30% of genes that had not been associated with another GO term or any other GO term with a more significant P-value (Ji and Tian, 2009).

In vitro kinase assays

In vitro kinase assays were performed according to Campbell (2014). After growing cultures to 0.5 OD, SLT2-3HA cultures were incubated at 37^{o}C for 1 hour, and HOG1-3HA cultures were grown in 0.8 M NaCl for 1 hour, cells were harvested and extracts prepared as described (Rosonina et al. 2009). Briefly, cells were agitated with glass beads in IP buffer for 30 minutes, then washed and pelleted. Cell extracts were IPed using HA antibody (ABM) and Sepharose-G beads at 4^{o}C for two hours, then washed three times with lysis buffer and twice with kinase buffer (25 mM Tris pH 7.5, 5 mM β-glycerophosphate, 2 mM DTT, 10 mM MgCl2, 10 mM NaF). Beads were then resuspended in 60 uL kinase buffer, with 200 nM GST-CTD (Hsin et al. 2014) and 10 mM ATP. Reaction mixtures were incubated for two hours at 30^{o}C with agitation every 10 minutes to keep beads suspended. 2X SDS sample buffer was then added and samples were boiled for 5 minutes before analysis by SDS PAGE and Western blot.

References

Akhtar MS, Heidemann M, Tietjen JR, Zhang DW, Chapman RD, Eick D, Ansari AZ. 2009. TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Mol Cell* **34:** 387-93.

Allen BL and Taatjes DJ. 2015. The Mediator complex: a central integrator of transcription. Nat Rev Mol Cell Biol 16: 155–166.

Allison LA, Moyle M, Shales M, Ingles CJ. 1985. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* **42**: 599-610.

Arias P, Díez-Muñiz S, García R, Nombela C, Rodríguez-Peña JM, Arroyo J. 2011. Genome-wide survey of yeast mutations leading to activation of the yeast cell integrity MAPK pathway: Novel insights into diverse MAPK outcomes. *BMC Genomics* **12**: 390.

Arigo JT, Eyler DE, Carroll KL, Corden JL. 2006. Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol Cell* **23**: 841–851.

Baetz K, Moffat J, Haynes J, Chang M, Andrews B. 2001. Transcriptional coregulation by the cell integrity mitogen-activated protein kinase Slt2 and the cell cycle regulator Swi4. *Mol Cell Biol* **21**: 6515-28.

Baillat D, Hakimi MA, Naar AM, Shilatifard A, Cooch N, Shiekhattar R. 2005. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**: 265-276.

Baskaran R, Dahmus ME, Wang JY. 1993. Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. *Proc Natl Acad Sci USA* **90**: 11167-71.

Bonnet F, Vigneron M, Bensaude O, Dubois MF. 1999. Transcription-independent phosphorylation of the RNA polymerase II C-terminal domain (CTD) involves ERK kinases (MEK1/2). *Nucleic Acids Res* **27**: 4399–4404.

Bose S, Dutko JA, Zitomer RS. 2005. Genetic factors that regulate the attenuation of the general stress response of yeast. *Genetics* **169:** 1215–1226.

Brewster JL, Gustin MC. 2014. Hog1: 20 years of discovery and impact. *Sci Signal* **7**: re7.

Burugula BB, Jeronimo C, Pathak R, Jones JW, Robert F, Govind CK. 2014. Histone deacetylases and phosphorylated polymerase II C-terminal domain recruit Spt6 for cotranscriptional histone reassembly. *Mol Cell Biol* **34**:4115-29.

Cai H, Kauffman S, Naider F, Becker JM. 2006. Genomewide screen reveals a wide regulatory network for di/tripeptide utilization in Saccharomyces cerevisiae. *Genetics* **172**: 1459-76.

Campbell, PA. 2014. IP-Kinase Assay. Bio-protocol 4: e1059.

Chapman RD, Heidemann M, Albert TK, Mailhammer R, Flatley A, Meisterernst M, Kremmer E, Eick D. 2007. Transcribing RNA Polymerase II is phosphorylated at CTD residue Serine-7. *Science* **318**: 1780-1782.

Chasman D, Ho YH, Berry DB, Nemec CM, MacGilvray ME, Hose J, Merrill AE, Lee MV, Will JL, Coon JJ, Ansari AZ, Craven M, Gasch AP. 2014. Pathway connectivity and signaling coordination in the yeast stress-activated signaling network. *Mol Syst Biol* **10**: 759.

Chen RE and Thorner J. 2007. Function and regulation in MAPK signaling pathways: lessons learned from the yeast Saccharomyces cerevisiae. *Biochim Biophys Acta* **1773:** 1311-40.

Chi A, Huttenhower C, Geer LY, Coon JJ, Syka JE, Bai DL, Shabanowitz J, Burke DJ, Troyanskaya OG, Hunt DF. 2007. Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. *Proc Natl Acad Sci USA* **104**: 2193-8.

Chi Y, Huddleston MJ, Zhang X, Young RA, Annan RS, Carr SA, Deshaies RJ. 2001. Negative regulation of Gcn4 and Msn2 transcription factors by Srb10 cyclin-dependent kinase. *Genes Dev* **15**: 1078-92.

Cho EJ, Takagi T, Moore CR, Buratowski S. 1997. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev* **11:** 3319-3326.

Cohen TJ, Lee K, Rutkowski LH, Strich R. 2003. Ask10p mediates the oxidative stress-induced destruction of the Saccharomyces cerevisiae C-type cyclin Une3p/Srb11p. *Eukaryot Cell* **2**: 962-970.

Conaway RC and Conaway JW. 2015. Orchestrating transcription with the pol II CTD. *Nat Rev Mol Cell Bio* **16**: 128.

Cooper KF, Mallory MJ, Smith JB, Strich R. 1997. Stress and developmental regulation of the yeast C-type cyclin Ume3p (Srb11p/Ssn8p). *EMBO J* **16:** 4665–4675.

Corden JL. 1990. Tails of RNA polymerase II. *Trends in Biochemical Sciences* **15**: 383-7.

Corden JL, Cadena DL, Ahearn JM Jr, Dahmus ME. 1985. A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc Natl Acad Sci U S A.* **82**: 7934-8.

Czudnochowski N, Bösken CA, Geyer M. 2012. Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. *Nat Comm* **3**: 842.

Descostes N, Heidemann M, Spinelli L, Schüller R, Maqbool MA, Fenouil R, Koch F, Innocenti C, Gut M, Gut I, Eick D, Andrau JC. 2014. Tyrosine phosphorylation of RNA polymerase II CTD is associated with antisense promoter transcription and active enhancers in mammalian cells. *Elife* **3**: e02105.

Devaiah BN, Lewis BA, Cherman N, Hewitt MC, Albrecht BK, Robey PG, Ozato K, Sims RJ 3rd, Singer DS. 2012. BRD4 is an atypical kinase that phosphorylates Serine2 of the RNA polymerase II carboxy-terminal domain. *Proc Natl Acad Sci USA* **109**: 6927-32.

De Virgilio C. 2012. The essence of yeast quiescence. FEMS Microbiol Rev. 36: 306-339

Dynlacht BD. 1997. Regulation of transcription by proteins that control the cell cycle. *Nature* **389**: 149-152.

Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S. 2007. Serine 7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* **318**: 1777-9.

Eick D and Geyer M. 2013. The RNA Polymerase II Carboxy-Terminal Domain (CTD) code. *Chem. Rev.* **113:** 8456–8490.

Feaver WJ, Svejstrup JQ, Henry NL, Kornberg RD. 1994. Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. *Cell* **79**: 1103-9.

Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* **11**: 4241-57.

Gene Ontology Consortium. 2015. Gene Ontology Consortium: going forward. *Nucleic Acids Res* **43** (Database issue): D1049-56.

Goshen-Lago T, Goldberg-Carp A, Melamed D, Darlyuk-Saadon I, Bai C, Ahn NG, Admon A, Engelberg D. 2016. Variants of the yeast MAPK Mpk1 are fully functional independently of activation loop phosphorylation. *Mol Biol Cell* **27**: 2771–2783.

Govind CK, Zhang F, Qiu H, Hofmeyer K, Hinnebusch AG. 2007. Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. *Mol Cell* **25**: 31-42.

Grossmann S, Bauer S, Robinson PN, Vingron M. 2007. Improved detection of overrepresentation of Gene-Ontology annotations with parent child analysis. *Bioinformatics* **23**: 3024-31.

Harlen KM, Trotta KL, Smith EE, Mosaheb MM, Fuchs SM, Churchman LS. 2016. Comprehensive RNA polymerase II interactomes reveal distinct and varied roles for each phospho-CTD residue. *Cell Rep* **15**: 2147-58.

Heidemann M and Eick D. 2012. Tyrosine-1 and threonine-4 phosphorylation marks complete the RNA polymerase II CTD phospho-code. *RNA Biol* **9:** 1144–1146.

Hengartner CJ, Thompson CM, Zhang J, Chao DM, Liao SM, Koleske AJ, Okamura S, Young RA. 1995. Association of an activator with an RNA polymerase II holoenzyme. *Genes Dev* **9**: 897-910.

Hintermair C, Heidemann M, Koch F, Descostes N, Gut M, Gut I, Fenouil R, Ferrier P, Flatley A, Kremmer E, Chapman RD, Andrau JC, Eick D. 2012. Threonine-4 of

mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. *EMBO J* **31**: 2784-97.

Hoekstra MF, Dhillon N, Carmel G, DeMaggio AJ, Lindberg RA, Hunter T, Kuret J. 1995. Budding and fission yeast casein kinase I isoforms have dual-specificity protein kinase activity. *Mol Biol Cell* **5**: 877-86.

Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, Park JY, Yehia G, Tian B. 2013. Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods* **10**: 133-139.

Hsin JP, Li W, Hoque M, Tian B, Manley JL. 2014. RNAP II CTD tyrosine 1 performs diverse functions in vertebrate cells. *eLife* **3**: e02112.

Hsin JP and Manley JL. 2012. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev* **26**: 2119-37.

Hsin JP, Sheth A, Manley JL. 2011. RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3’ end processing. *Science* **334**: 683-686.

Irqeba AA, Li Y, Panahi M, Zhu M, Wang Y. 2014. Regulating global sumoylation by a MAP kinase Hog1 and its potential role in osmo-tolerance in yeast. *PLoS One* **9**: e87306.

Jasnovidova O, Krejcikova M, Kubicek K, Stefl R. 2017. Structural insight into recognition of phosphorylated threonine-4 of RNA polymerase II C-terminal domain by Rtt103p. *EMBO Rep* **18**: 906-913.

Ji Z and Tian B. 2009. Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. *PLoS One* **4**: e8419.

Jin C, Strich R, Cooper KF. 2014. Slt2p phosphorylation induces cyclin C nuclear-to-cytoplasmic translocation in response to oxidative stress. *Molecular Biology of the Cell* **25**: 1396-1407.

Jung US, Sobering AK, Romeo MJ, Levin DE. 2002. Regulation of the yeast Rlm1 transcription factor by the Mpk1 cell wall integrity MAP kinase. *Mol Microbiol* **46**: 781–789.

Kahana-Edwin S, Stark M, Kassir Y. 2013. Multiple MAPK cascades regulate the transcription of IME1, the master transcriptional activator of meiosis in *Saccharomyces cerevisiae*. *PLoS One* **8**: e78920.

Kamada Y, Jung US, Piotrowski J, Levin DE. 1995. The protein kinase C-activated MAP kinase pathway of Saccharomyces cerevisiae mediates a novel aspect of the heat shock response. *Genes Dev* **9**: 1559-71.

Kassis S, Melhuish T, Annan RS, Chen SL, Lee JC, Livi GP, Creasy CL. 2000. Saccharomyces cerevisiae Yak1p protein kinase autophosphorylates on tyrosine residues and phosphorylates myelin basic protein on a C-terminal serine residue. *Biochem J* **348**: 263-72.

Kim KY and Levin DE. 2010. Transcriptional reporters for genes activated by cell wall stress through a non-catalytic mechanism involving Mpk1 and SBF. *Yeast* **27**: 541–548.

Kim KY and Levin DE. 2011. Mpk1 MAPK association with the Paf1 complex blocks Sen1-mediated premature transcription termination. *Cell* **144:** 745-56.

Kim KY, Truman AW, Caesar S, Schlenstedt G, Levin DE. 2010. Yeast Mpk1 cell wall integrity mitogen-activated protein kinase regulates nucleocytoplasmic shuttling of the Swi6 transcriptional regulator. *Mol Biol Cell* **21**: 1609–1619.

Kim KY, Truman AW, Levin DE. 2008. Yeast Mpk1 mitogen-activated protein kinase activates transcription through Swi4/Swi6 by a noncatalytic mechanism that requires upstream signal. *Mol Cell Biol* **28**: 2579-89.

Kim M, Krogan NJ, Vasiljeva L, Rando OJ, Nedea E, Greenblatt JF, Buratowski S. 2004. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432:** 517-22.

Koch C, Schleiffer A, Ammerer G, Nasmyth K. 1996. Switching transcription on and off during the yeast cell cycle: Cln/Cdc28 kinases activate bound transcription factor SBF (Swi4/Swi6) at start, whereas Clb/Cdc28 kinases displace it from the promoter in G2. *Genes Dev* **10**: 129-41.

Koutelou E, Hirsch CL, Dent SY. 2010. Multiple faces of the SAGA complex. *Curr Opin Cell Biol* **22**: 374-82.

Krasley E, Cooper KF, Mallory MJ, Dunbrack R, Strich R. 2006. Regulation of the oxidative stress response through Slt2p-dependent destruction of cyclin C in Saccharomyces cerevisiae. *Genetics* **172:** 1477-86.

Kuranda K, Leberre V, Sokol S, Palamarczyk G, François J. 2006. Investigating the caffeine effects in the yeast Saccharomyces cerevisiae brings new insights into the connection between TOR, PKC and Ras/cAMP signalling pathways. *Molecular Microbiology* **61**: 1147–1166.

Lee K, Du C, Horn M, Rabinow L. 1996. Activity and autophosphorylation of LAMMER protein kinases. *J Biol Chem* **271**: 27299-303.

Lee KK, Florens L, Swanson SK, Washburn MP, Workman JL. 2005. The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. *Mol Cell Biol* **25**: 1173-82.

Levin DE. 2005. Cell wall integrity signaling in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev* **69**: 262–291.

Levin DE. 2011. Regulation of cell wall biogenesis in Saccharomyces cerevisiae: the cell wall integrity signaling pathway. *Genetics* **189**: 1145–1175.

Levin-Salomon V, Maayan I, Avrahami-Moyal L, Marbach I, Livnah O, Engelberg D. 2009. When expressed in yeast, mammalian mitogen-activated protein kinases lose proper regulation and become spontaneously phosphorylated. *Biochem J* **417**: 331-40.

Liu P, Kenney JM, Stiller JW, Greenleaf AL. 2010. Genetic organization, length conservation, and evolution of RNA polymerase II carboxyl-terminal domain. *Mol Biol Evol* **27**: 2628–2641.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Lunde BM, Reichow SL, Kim M, Suh H, Leeper TC, Yang F, Mutschler H, Buratowski S, Meinhart A, Varani G. 2010. Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **17**: 1195-201.

Luo W, Johnson AW, Bentley DL. 2006. The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes Dev* **20**: 954-65.

Maayan I, Beenstock J, Marbach I, Tabachnick S, Livnah O, Engelberg D. 2012. Osmostress induces autophosphorylation of Hog1 via a C-terminal regulatory region that is conserved in p38α. *PLoS One* **7**: e44749.

Madden K, Sheu YJ, Baetz K, Andrews B, Snyder M. 1997. SBF cell cycle regulator as a target of the yeast PKC-MAP kinase pathway. *Science* **275**: 1781-4.

Madrid M, Soto T, Khong HK, Franco A, Vicente J, Pérez P, Gacto M, Cansado J. 2006. Stress-induced response, localization, and regulation of the Pmk1 cell integrity pathway in Schizosaccharomyces pombe. *J Biol Chem* **281**: 2033-43.

Martín H, Rodríguez-Pachón JM, Ruiz C, Nombela C, Molina M. 2000. Regulatory mechanisms for modulation of signaling through the cell integrity Slt2-mediated pathway in *Saccharomyces cerevisiae*. *J Biol Chem* **275**: 1511-1519.

Mayer A, Lidschreiber M, Siebert M, Leike K, Söding J, Cramer P. 2010. Uniform transitions of the general RNA polymerase II transcription complex. *Nat Struct Mol Bio* **17**: 1272–1278.

Mayer A, Heidemann M, Lidschreiber M, Schreieck A, Sun M, Hintermair C, Kremmer E, Eick D, Cramer P. 2012. CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. *Science* **336**: 1723-5.

Mikula M, Skrzypczak M, Goryca K, Paczkowska K, Ledwon JK, Statkiewicz M, Kulecka M, Grzelak M, Dabrowska M, Kuklinska U, Karczmarski J, Rumienczyk I, Jastrzebski K, Miaczynska M, Ginalski K, Bomsztyk K, Ostrowski J. 2016. Genome-wide co-localization of active EGFR and downstream ERK pathway kinases mirrors mitogen-inducible RNA polymerase 2 genomic occupancy. *Nucleic Acids Res* **44**: 10150-10164.

Mizunuma M, Hirata D, Miyahara K, Tsuchiya E, Miyakawa T. 1998. Role of calcineurin and Mpk1 in regulating the onset of mitosis in budding yeast. *Nature* **392**: 303-306.

Modesti A, Bini L, Carraresi L, Magherini F, Liberatori S, Pallini V, Manao G, Pinna LA, Raugei G, Ramponi G. 2001. Expression of the small tyrosine phosphatase (Stp1) in Saccharomyces cerevisiae: a study on protein tyrosine phosphorylation. *Electrophoresis* **22**: 576–585.

Mollapour M, Tsutsumi S, Donnelly AC, Beebe K, Tokita MJ, Lee MJ, Lee S, Morra G, Bourboulia D, Scroggins BT, Colombo G, Blagg BS, Panaretou B, Stetler-Stevenson WG, Trepel JB, Piper PW, Prodromou C, Pearl LH, Neckers L. 2010. Swe1Wee1-dependent tyrosine phosphorylation of Hsp90 regulates distinct facets of chaperone function. *Mol Cell* **37**: 333-43.

Moreno-Torres M, Jaquenoud M, De Virgilio C. 2015. TORC1 controls G1–S cell cycle transition in yeast via Mpk1 and the greatwall kinase pathway. *Nat Commun* **6**: 8256.

Napolitano G, Lania L, Majello B. 2014. RNA polymerase II CTD modifications: how many tales from a single tail. *J Cell Physiol* **229**: 538-44.

Nelson C, Goto S, Lund K, Hung W, Sadowski I. 2003. Srb10/Cdk8 regulates yeast filamentous growth by phosphorylating the transcription factor Ste12. *Nature* **421**: 187-190.

Ng HH, Robert F, Young RA, Struhl K. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11:** 709-19.

O'Duibhir E, Lijnzaad P, Benschop JJ, Lenstra TL, van Leenen D, Groot Koerkamp MJ, Margaritis T, Brok MO, Kemmeren P, Holstege FC. 2014. Cell cycle population effects in perturbation studies. *Mol Syst Biol* **10**: 732.

Ohkuni K and Yamashita I. 2000. A transcriptional autoregulatory loop for KIN28-CCL1 and SRB10-SRB11, each encoding RNA polymerase II CTD kinase-cyclin pair, stimulates the meiotic development of S. cerevisiae. *Yeast* **16**: 829-46.

O'Reilly D, Kuznetsova OV, Laitem C, Zaborowska J, Dienstbier M, Murphy S. 2014. Human snRNA genes use polyadenylation factors to promote efficient transcription termination. *Nucleic Acids Res* **42**: 264-75.

Pearson G, Robinson F, Gibson TB, Xu B, Karandikar M, Berman K, Cobb MH. 2011. Mitogen-Activated Protein (MAP) kinase pathways: Regulation and physiological functions. *Endocrine Reviews* **22**: 153–83.

Phatnani HP and Greenleaf AL. 2006. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev* **20**: 2922-2936.

Pokholok DK, Zeitlinger J, Hannett NM, Reynolds DB, Young RA. 2006. Activated signal transduction kinases frequently occupy target genes. *Science* **313**: 533–536.

Porrua O and Libri D. 2015. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat Rev Mol Cell Biol* **16**: 190–202.

Posas F and Saito H. 1997. Osmotic activation of the HOG MAPK pathway via Ste11p MAPKKK: scaffold role of Pbs2p MAPKK. *Science* **276**: 1702–1705.

Posas F, Takekawa M, Saito H. 1998. Signal transduction by MAP kinase cascades in budding yeast. *Curr Op Mol Biol* **1** :175–182.

Poschke H, Dees M, Chang M, Amberkar S, Kaderali L, Rothstein R, Luke B. 2012. Rif2 promotes a telomere fold-back structure through Rpd3L recruitment in budding yeast. *PLoS Genet* **8**: e1002960.

Pujol-Carrion N1, Petkova MI, Serrano L, de la Torre-Ruiz MA. 2013. The MAP kinase Slt2 is involved in vacuolar function and actin remodeling in Saccharomyces cerevisiae mutants affected by endogenous oxidative stress. *Appl Environ Microbiol* **79**: 6459-71.

Rep M, Krantz M, Thevelein JM, Hohmann S. 2000. The transcriptional response of Saccharomyces cerevisiae to osmotic shock. Hot1p and Msn2p/Msn4p are

required for the induction of subsets of high osmolarity glycerol pathway-dependent genes. *J Biol Chem* **275**: 8290-300.

Robinson PJ, Trnka MJ, Bushnell DA, Davis RE, Mattei PJ, Burlingame AL, Kornberg RD. 2016. Structure of a complete Mediator-RNA Polymerase II Pre-Initiation Complex. *Cell* **166**: 1411-1422.

Rosonina E, Duncan SM & Manley JL. 2010. SUMO functions in constitutive transcription and during activation of inducible genes in yeast. *Genes Dev* **24**: 1242-1252.

Rosonina E, Duncan SM & Manley JL. 2012. Sumoylation of transcription factor Gcn4 facilitates its Srb10-mediated clearance from promoters in yeast. *Genes Dev* **26**: 350-355.

Rosonina E, Willis IM, Manley JL. 2009. Sub1 functions in osmoregulation and in transcription by both RNA polymerases II and III. *Mol Cell Biol* **29**: 2308-2321.

Rosonina E, Yurko N, Li W, Hoque M, Tian B, Manley JL. 2014. Threonine-4 of the budding yeast RNAP II CTD couples transcription with Htz1-mediated chromatin remodeling. *Proc Natl Acad Sci USA* **111**: 11924-31.

Satoh R, Morita T, Takada H, Kita A, Ishiwata S, Doi A, Hagihara K, Taga A, Matsumura Y, Tohda H, Sugiura R. 2009. Role of the RNA-binding protein Nrd1 and Pmk1 mitogen-activated protein kinase in the regulation of myosin mRNA stability in fission yeast. *Mol Biol Cell* **20**: 2473–2485.

Schüller R, Forné I, Straub T, Schreieck A, Texier Y, Shah N, Decker TM, Cramer P, Imhof A, Eick D. 2016. Heptad-specific phosphorylation of RNA Polymerase II CTD. *Mol Cell* **61**: 305-14.

Schwer B, Bitton DA, Sanchez AM, Bähler J, Shuman S. 2014. Individual letters of the RNA polymerase II CTD code govern distinct gene expression programs in fission yeast. *PNAS* **111**: 4185–4190.

Schwer B, Sanchez AM, Shuman S. 2012. Punctuation and syntax of the RNA polymerase II CTD code in fission yeast. *Proc Natl Acad Sci* **109**: 18024–18029.

Schwer B, Shuman S. 2011. Deciphering the RNA polymerase II CTD code in fission yeast. *Mol Cell* **43**: 311–318.

Smorodinsky-Atias K, Goshen-Lago T, Goldberg-Carp A, Melamed D, Shir A, Mooshayef N, Beenstock J, Karamansha Y, Darlyuk-Saadon I, Livnah O, Ahn NG, Admon A, Engelberg D. 2016. Intrinsically active variants of Erk oncogenically transform cells and disclose unexpected autophosphorylation capability that is independent of TEY phosphorylation. *Mol Biol Cell* **27**: 1026–1039.

Soriano-Carot M, Bañó MC, Igual JC. 2012. The yeast mitogen-activated protein kinase Slt2 is involved in the cellular response to genotoxic stress. *Cell Division* **7**: 1.

Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA. 2006. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol Cell* **24**: 735-46.

Strich R and Cooper KF. 2014. The dual role of cyclin C connects stress regulated gene expression to mitochondrial dynamics. *Microbial Cell* **1**: 318–324.

Suh H, Ficarro SB, Kang UB, Chun Y, Marto JA, Buratowski S. 2016. Direct analysis of phosphorylation sites on the Rpb1 C-terminal domain of RNA polymerase II. *Mol Cell* **61**: 297-304.

Toda T, Dhut S, Superti-Furga G, Gotoh Y, Nishida E, Sugiura R, Kuno T. 1996. The fission yeast pmk1+ gene encodes a novel mitogen-activated protein kinase homolog which regulates cell integrity and functions coordinately with the protein kinase C pathway. *Mol Cell Biol* **16**: 6752-6764

Tong AH and Boone C. 2006. Synthetic genetic array analysis in Saccharomyces cerevisiae. *Methods Mol Biol* **313**: 171-92.

Torres L, Martín H, García-Saez MI, Arroyo J, Molina M, Sánchez M, Nombela C. 1991. A protein kinase gene complements the lytic phenotype of Saccharomyces cerevisiae lyt2 mutants. *Mol Microbiol* **5**: 2845-54.

Trigon S, Serizawa H, Conaway JW, Conaway RC, Jackson SP, Morange M. 1998. Characterization of the residues phosphorylated in vitro by different C-terminal domain kinases. *J Biol Chem* **273**: 6769–6775.

Truman AW, Kim KY, Levin DE. 2009. Mechanism of Mpk1 Mitogen-Activated Protein Kinase binding to the Swi4 transcription factor and its regulation by a novel caffeine-induced phosphorylation. *Mol Cell Biol* **29**: 6449–6461.

Tyner C, Barber GP, Casper J, Clawson H, Diekhans M, Eisenhart C, Fischer CM, Gibson D, Gonzalez JN, Guruvadoo L, Haeussler M, Heitner S, Hinrichs AS, Karolchik D, Lee BT, Lee CM, Nejad P, Raney BJ, Rosenbloom KR, Speir ML, Villarreal C, Vivian J, Zweig AS, Haussler D, Kuhn RM, Kent WJ. 2017. The UCSC Genome Browser database: 2017 update. *Nucleic Acids Res* **45** (Database issue): D626-D634.

van de Peppel J, Kettelarij N, van Bakel H, Kockelkorn TT, van Leenen D, Holstege FC. 2005. Mediator expression profiling epistasis reveals a signal transduction pathway with antagonistic submodules and highly specific downstream targets. *Mol Cell* **19**: 511-22.

van Voorst F, Houghton-Larsen J, Jønson L, Kielland-Brandt MC, Brandt A. 2006. Genome-wide identification of genes required for growth of Saccharomyces cerevisiae under ethanol stress. *Yeast* **23**: 351-9.

Vasiljeva L, Kim M, Mutschler H, Buratowski S, Meinhart A. 2008. The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **15:** 795-804.

Verghese J, Abrams J, Wang Y, Morano KA. 2012. Biology of the heat shock response and protein chaperones: budding yeast (*Saccharomyces cerevisiae*) as a model system. *Microbiol Mol Biol Rev* **76**: 115-158.

West ML and Corden JL. 1995. Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. *Genetics* **140**: 1223-33.

Wilson MA, Koutelou E, Hirsch C, Akdemir K, Schibler A, Barton MC, Dent SY. 2011. Ubp8 and SAGA regulate Snf1 AMP kinase activity. *Mol Cell Biol* **31**: 3126-35.

Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM. 2009. Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**: 1033-7.

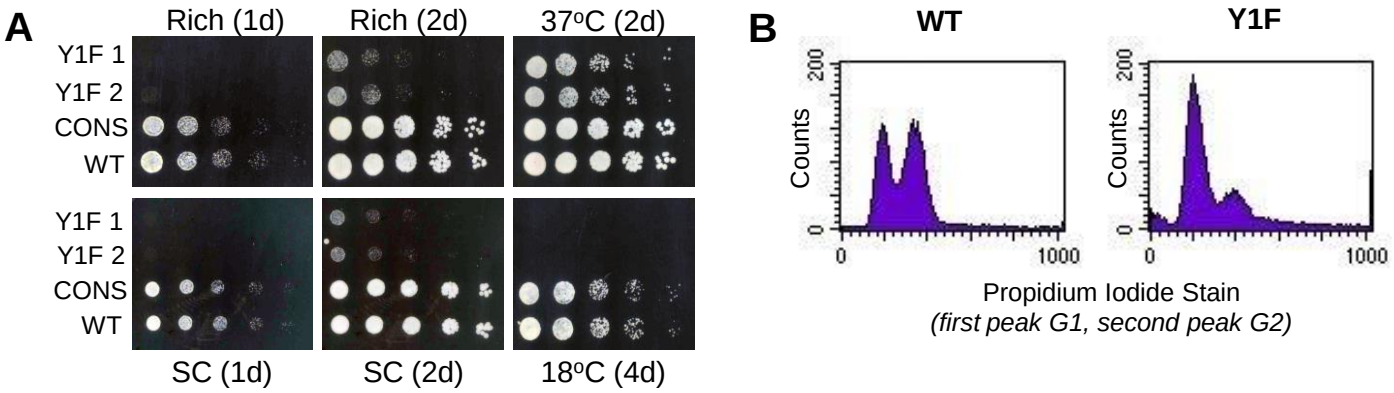
Yang SH, Sharrocks AD, Whitmarsh AJ. 2013. MAP kinase signalling cascades and transcriptional regulation. *Gene* **513**: 1-13.

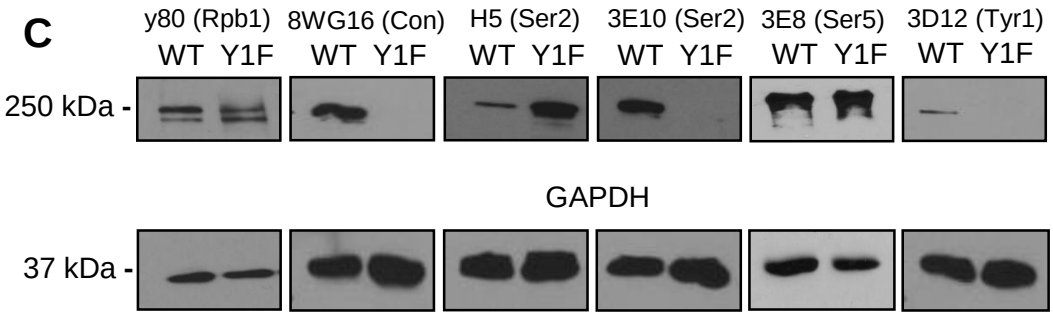
Yurko NM and Manley JL. 2017. The RNA polymerase II CTD “orphan” residues: Emerging insights into the functions of Tyr-1, Thr-4 and Ser-7. *Transcription*, in press.

Yuryev A and Corden JL. 1996. Suppression analysis reveals a functional difference between the serines in positions two and five in the consensus sequence of the C-terminal domain of yeast RNA polymerase II. *Genetics* **143**: 661-71.

Zhang H and Siede W. 2004. Analysis of the budding yeast Saccharomyces cerevisiae cell cycle by morphological criteria and flow cytometry. *Methods Mol Biol* **241**: 77-91.

Figure 1





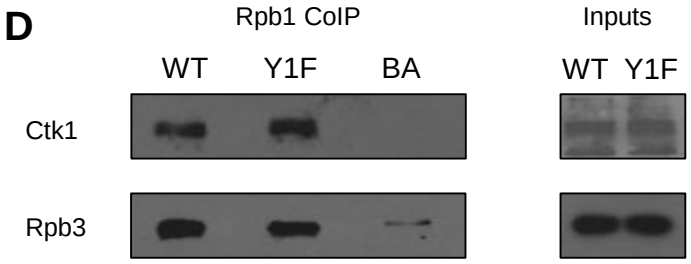


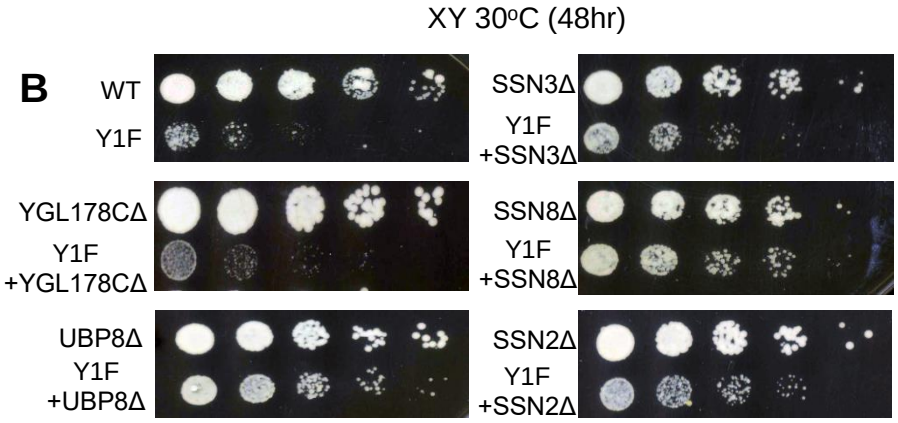
Figure 1: RNAP II CTD Tyr to Phe substitution results in severe growth defects. (A) Tetrad dissection of heterozygous diploid Y1F CTD strain (Y1F). Y1F CTD in haploid tetrads marked with NAT resistance and is present in all small-size tetrads. Yeast spot assay comparing growth of WT, CONS, and two Y1F clones, with serial fivefold dilutions, on rich or synthetic complete (SC) medium at 30oC, or on rich medium at indicated temperature, for indicated number of days. (B) Cell-cycle assay showing unsynchronized cells stained with propidium iodide, the two peaks representing cells in G1/G0 and G2, respectively. (C) Western blot analysis of cell lysates derived from WT and Y1F strains using antibodies recognizing the N-terminus of Rpb1 (y80), unphosphorylated consensus repeats (8WG16), Ser2P (3E10 and H5), Ser5P (3E8), Tyr1P (3D12) and GAPDH as indicated. (D) Western blot analysis of an Rpb1 co-IP (using y-80 antibody), probing for Ctk1 interaction (using 3HA tagged Ctk1 and HA antibody) with RNAP II (normalized to Rpb3). 5% inputs shown. All results shown are representative of three independent experiments.

Figure 2

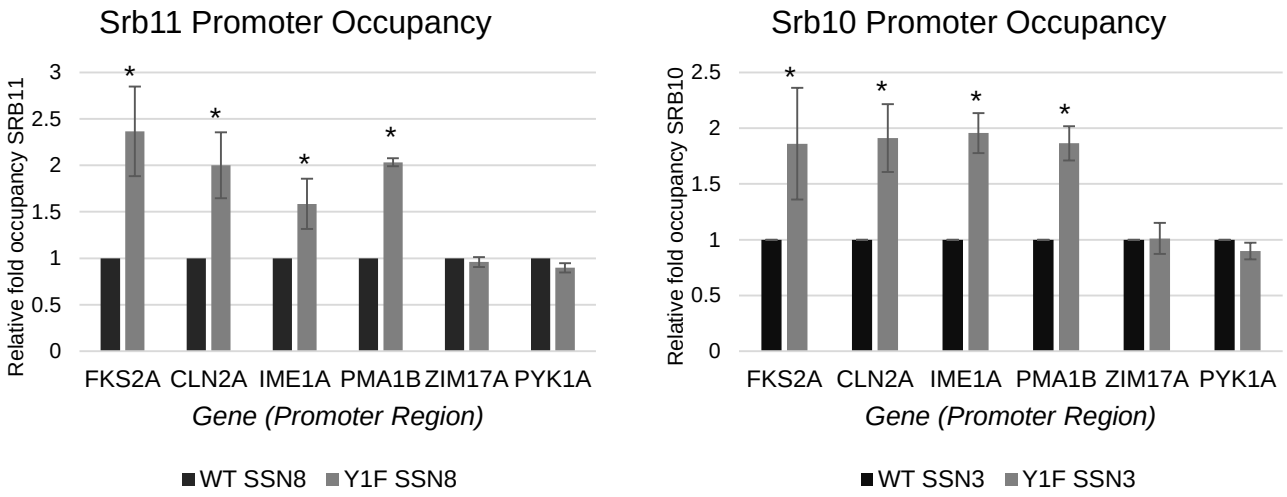
A

Suppressor Screen Rescues	
SSN2	Subunit of Mediator Complex; Srb9
SSN8	Subunit of Mediator Complex (cyclin); Srb11
SSN3	Subunit of Mediator Complex (kinase); Srb10
SPT2	Interacts with Histones/Swi-Snf
UBP8	SAGA Component
HRP1	Subunit of Cleavage Factor I

B



C



*: $p < 0.05$

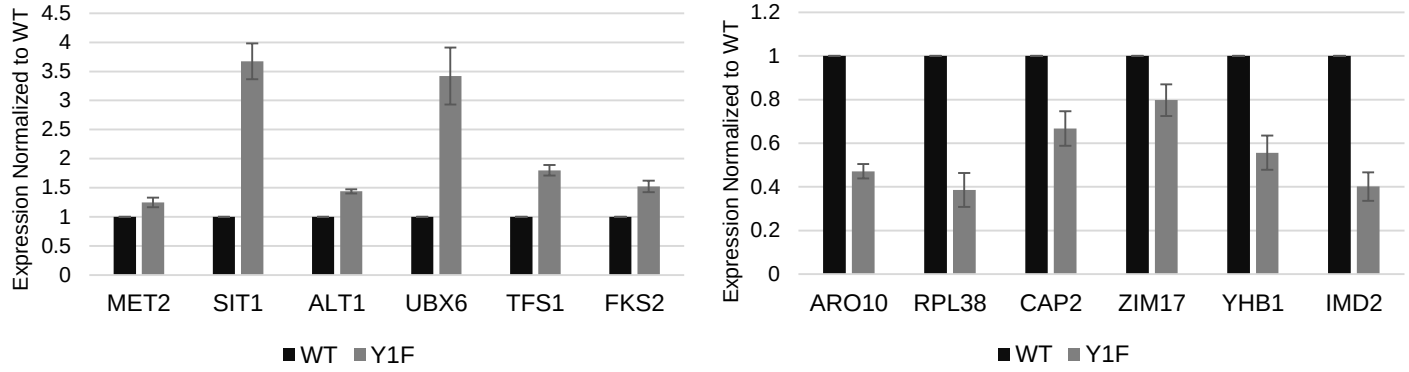
Figure 2: SGA analysis identifies suppressors of Y1F growth defects. (A) List of genes showing synthetic rescue of Y1F strain in two SGA suppressor screens. Genes displaying strongest interactions and involved in RNAP II transcription are shown (refer to Dataset S1 for a complete list). (B) Confirmation of genetic interaction between Y1F and SRB9/10/11 and UBP8. Strains of indicated genotypes were generated by recombinant transformation and growth compared by spot assay on rich medium with serial five-fold dilutions. YGL178C showed no interaction with Y1F in SGA screen and serves as a control. (C) ChIP analysis of Srb10/11 occupancy at promoter regions of indicated genes, using 3HA-tagged Srb10/11. Data normalized to WT signal at each gene; P values less than 0.05 are indicated (*). ChIP data is represented as mean +/- SE of three independent experiments.

Figure 3

A

Category	GO_ID	Upregulated GO terms (306 genes)	-log10(P-value)
BP	GO:0055072	iron ion homeostasis	3.6
BP	GO:0000041	transition metal ion transport	3.6
BP	GO:0006811	ion transport	3.0
CC	GO:0005618	cell wall	2.9
CC	GO:0071944	cell periphery	2.9
CC	GO:0000407	pre-autophagosomal structure	2.3
MF	GO:0046915	transition metal ion transmembrane transporter activity	2.9
MF	GO:0016667	oxidoreductase activity, acting on a sulfur group of donors	2.9
MF	GO:0051287	NAD binding	2.3
Category	**GO_ID**	**Downregulated GO terms (324 genes)**	**-log10(P-value)**
BP	GO:0043603	cellular amide metabolic process	7.5
BP	GO:1901564	organonitrogen compound metabolic process	6.0
BP	GO:0044271	cellular nitrogen compound biosynthetic process	5.8
CC	GO:0032991	macromolecular complex	6.0
CC	GO:0044444	cytoplasmic part	2.4
CC	GO:0044459	plasma membrane part	1.1
MF	GO:0005198	structural molecule activity	6.9
MF	GO:0032403	protein complex binding	2.0
MF	GO:0016903	oxidoreductase activity, acting on the aldehyde or oxo group of donors	1.6

B



C

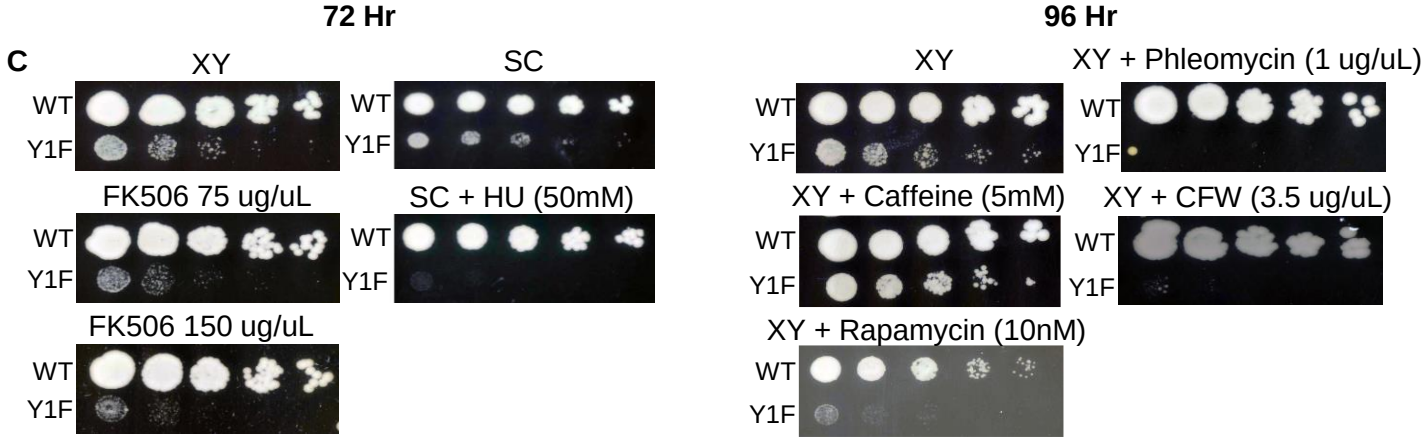


Figure 3: Expression of multiple genes is altered in Y1F. (A) List of enriched GO terms through Gene Ontology Analysis. The numbers of significantly altered genes used for GO Analysis were 306 up-regulated, 324 down-regulated. Dataset S2 presents a complete list of expression changes for 5,925 genes. (B) RT-PCR analysis confirming changes in RNA levels for indicated genes between WT and Y1F strains. (IMD2 not detected in original RNA-seq analysis). (C) Spot assays comparing growth of WT and Y1F strains, using serial five-fold dilutions, on rich and synthetic media with media containing indicated stress-inducing compounds.

Figure 4

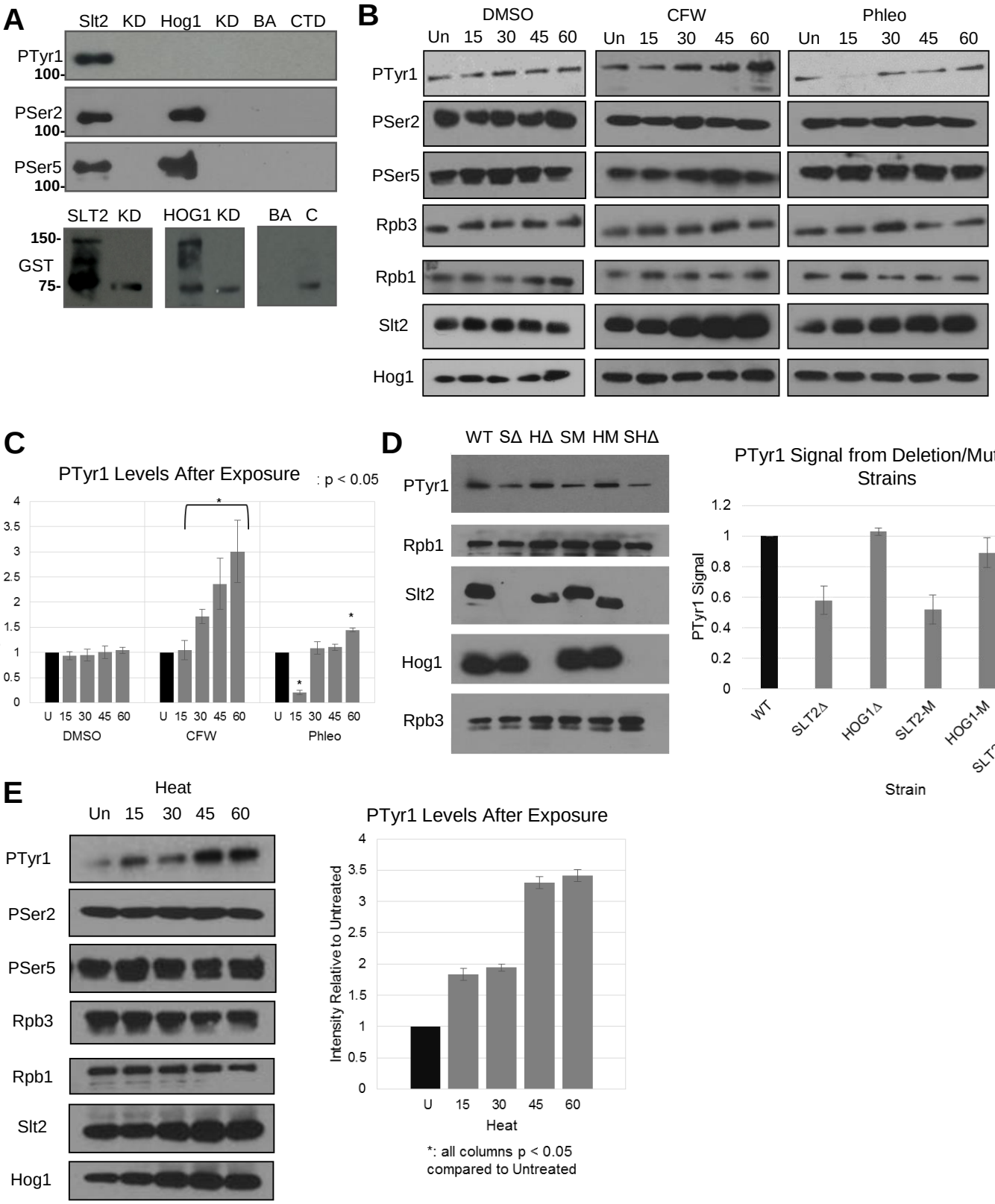

A
Slt2 KD Hog1 KD BA CTD
PTyr1
100-
PSer2
100-
PSer5
100-
SLT2 KD
HOG1 KD
BA C
150-
GST
75-
B
DMSO
CFW
Phleo
Un 15 30 45 60
PTyr1
PSer2
PSer5
Rpb3
Rpb1
Slt2
Hog1
C
PTyr1 Levels After Exposure
: p < 0.05
Intensity Relative to Untreated
U 15 30 45 60
DMSO
CFW
Phleo
D
WT SΔ HΔ SM HM SHΔ
PTyr1
Rpb1
Slt2
Hog1
Rpb3
PTyr1 Signal from Deletion/Mutation Strains
PTyr1 Signal
WT
SLT2Δ
HOG1Δ
SLT2-M
HOG1-M
SLT2Δ/HOG1Δ
Strain
E
Heat
Un 15 30 45 60
PTyr1
PSer2
PSer5
Rpb3
Rpb1
Slt2
Hog1
PTyr1 Levels After Exposure
Intensity Relative to Untreated
U 15 30 45 60
Heat
*: all columns p < 0.05
compared to Untreated

Figure 4 (continued)

Figure 4: MAP Kinase Slt2 phosphorylates Tyr1 in vitro and in vivo. (A) In vitro kinase assay using 3HA-tagged Slt2 and Hog1 extracted and IPed from activated cells. Following incubation with GST-CTD, proteins were resolved by SDS-PAGE and blots probed with the indicated antibodies. Lanes are Slt2-HA (SLT2) and kinase-dead (KD), Hog1-HA (HOG1) and kinase dead (KD), beads and antibody control (BA) and GST-CTD alone (CTD). (B) Western blot of Tyr1P levels after stress induction. DMSO, calcofluor white (CFW) and phleomycin (Phl) are shown. Rpb1, Slt2, Hog1, and Rpb3 were also probed using their respective antibodies. Time points are untreated (Un), 15, 30, 45 and 60 minutes. (C) Quantification of (B). Signals were normalized individually to Rpb3 levels, then collectively to uninduced control. (D) Western blot and quantification of Tyr1P levels in isogenic deletion strains. Protein extracts from strains with either SLT2/HOG1 deletions (SΔ/HΔ; double, SHΔ) or kinase-dead mutations (SM/HM) were blotted using 3D12 and normalized to WT. (E) Western blot of Tyr1P levels after heat stress (37oC); signals normalized as in (B). All results shown are representative of three independent experiments; data are represented as mean +/- SE of three independent experiments.

Figure 5

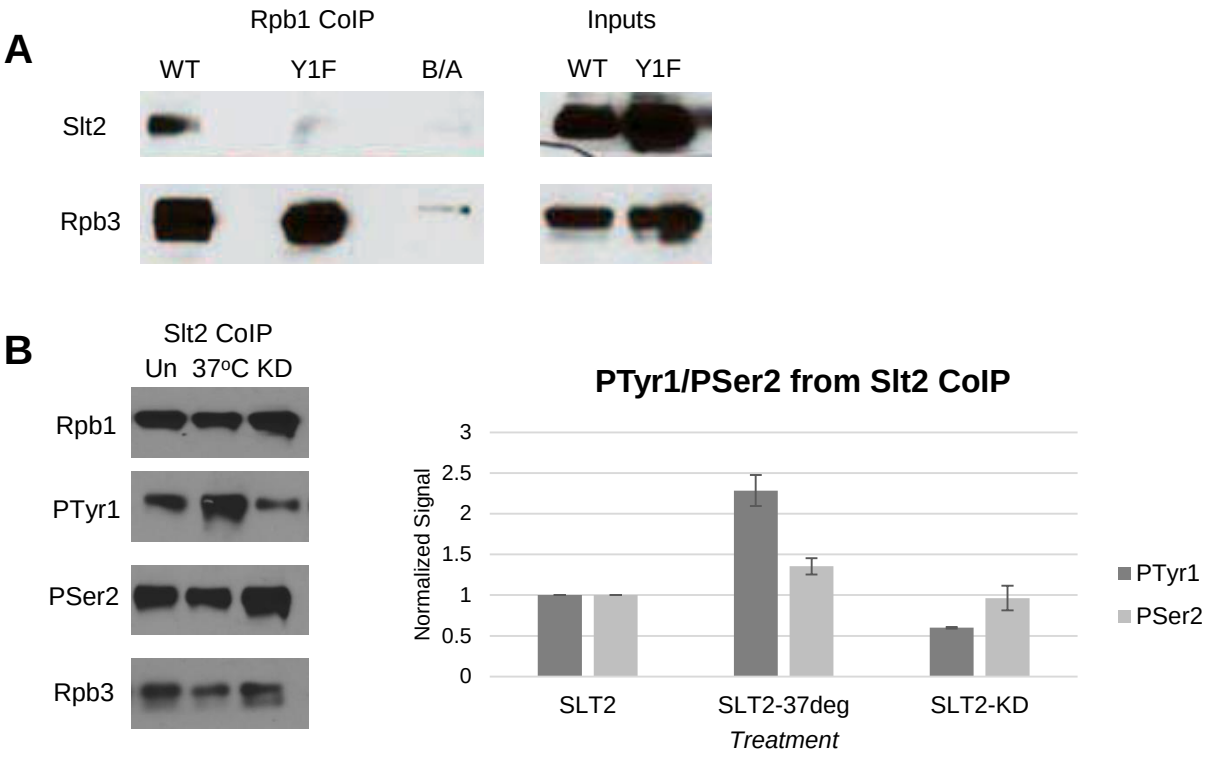


Figure 5: Differential association of Slt2 with WT and Y1F RNAP II. (A) Western blot analysis of a Rpb1 co-IP (using y-80 antibody), probing for Slt2 interaction with RNAP II (normalized to Rpb3). 5% inputs shown. (B) Western blot analysis of a Slt2-3HA co-IP, probing for total Rpb1 (y80), Tyr1-P Rpb1 (3D12) and Ser2-P Rpb1 (3E10), using Rpb3 as control for normalization. Untreated (Un), heat treated (37oC) and kinase dead (KD) shown. All results shown are representative of three independent experiments.

Figure 6

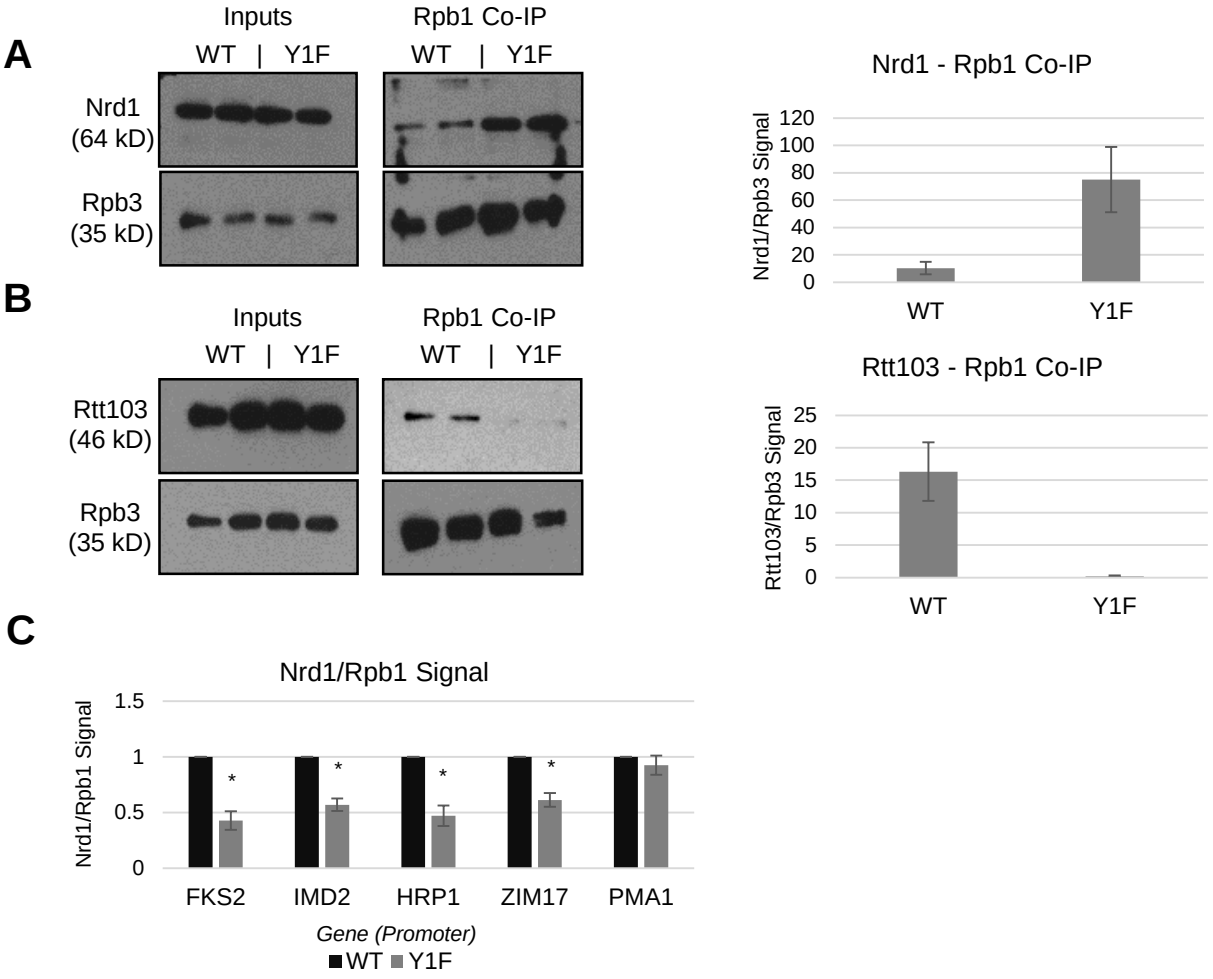

A
Inputs
Rpb1 Co-IP
WT | Y1F
WT | Y1F
Nrd1
(64 kD)
Rpb3
(35 kD)
Nrd1 - Rpb1 Co-IP
Nrd1/Rpb3 Signal
120
100
80
60
40
20
0
WT
Y1F
B
Inputs
Rpb1 Co-IP
WT | Y1F
WT | Y1F
Rtt103
(46 kD)
Rpb3
(35 kD)
Rtt103 - Rpb1 Co-IP
Rtt103/Rpb3 Signal
25
20
15
10
5
0
WT
Y1F
C
Nrd1/Rpb1 Signal
Nrd1/Rpb1 Signal
1.5
1
0.5
0
FKS2
IMD2
HRP1
ZIM17
PMA1
Gene (Promoter)
WT Y1F
D
Change in Nrd1 Occupancy
Nrd1 Occupancy
3
2.5
2
1.5
1
0.5
0
WT
Y1F
FKS2 Promoter
Change in Nrd1 Occupancy
Nrd1 Occupancy
WT
Y1F
PMA1 Promoter
Untreated
Heat
E
Change in Slt2 Occupancy
Slt2 Occupancy
2
1.5
1
0.5
0
WT
Y1F
FKS2 Promoter
Change in Slt2 Occupancy
Slt2 Occupancy
WT
Y1F
PMA1 Promoter
Untreate
d
Heat

Figure 6 (continued)

Figure 6: RNAP II Tyr1 and Slt2 affect Nrd1 recruitment to RNAP II and chromatin. (A) Western blot analysis of Nrd1 levels after Rpb1 co-IP (using y-80 antibody), with quantification. Nrd1 interaction with RNAP II quantified using Rpb3. (B) Western blot analysis of Rtt103 levels after Rpb1 co-IP (using y-80 antibody), with quantification. Rtt103 interaction with RNAP II quantified using Rpb3. (C) ChIP analysis of Nrd1 occupancy at promoter regions of indicated genes in WT and Y1F strains. (D) ChIP analysis of Nrd1 occupancy at promoter and body regions of model gene FKS2, before and after heat stress. (E) ChIP analysis of Slt2 occupancy at promoter and body regions of model gene FKS2, before and after heat stress. All data are represented as mean +/- SE of three independent experiments.

Figure 7

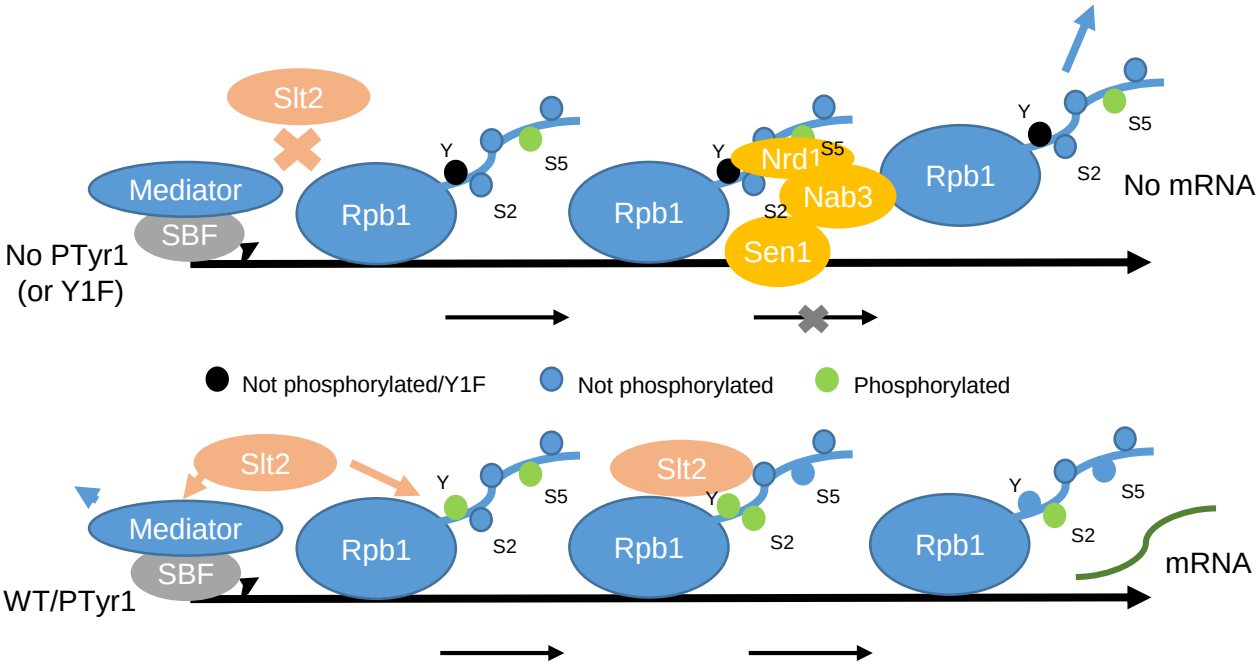


Figure 7: Role of CTD Tyr1 and Slt2 kinase in activating transcription of stress-inducible genes. In the absence of activation by a stress-induced signaling cascade, Slt2 is inactive and Tyr1 is not phosphorylated. The Nrd1-containing NNS complex is thus recruited to a primarily Ser5P CTD and transcripts are prematurely terminated and degraded. Activation of Slt2 enables the kinase to associate with RNAP II and Mediator/SBF, promoting Srb10/11 degradation and RNAP II Tyr1 phosphorylation. This combination of events facilitates activation and prevents termination factor (e.g., NNS) association with the CTD until the 3' end, when Tyr1 is dephosphorylated, the transcript 3' processed and transcription terminated.

Figure S1

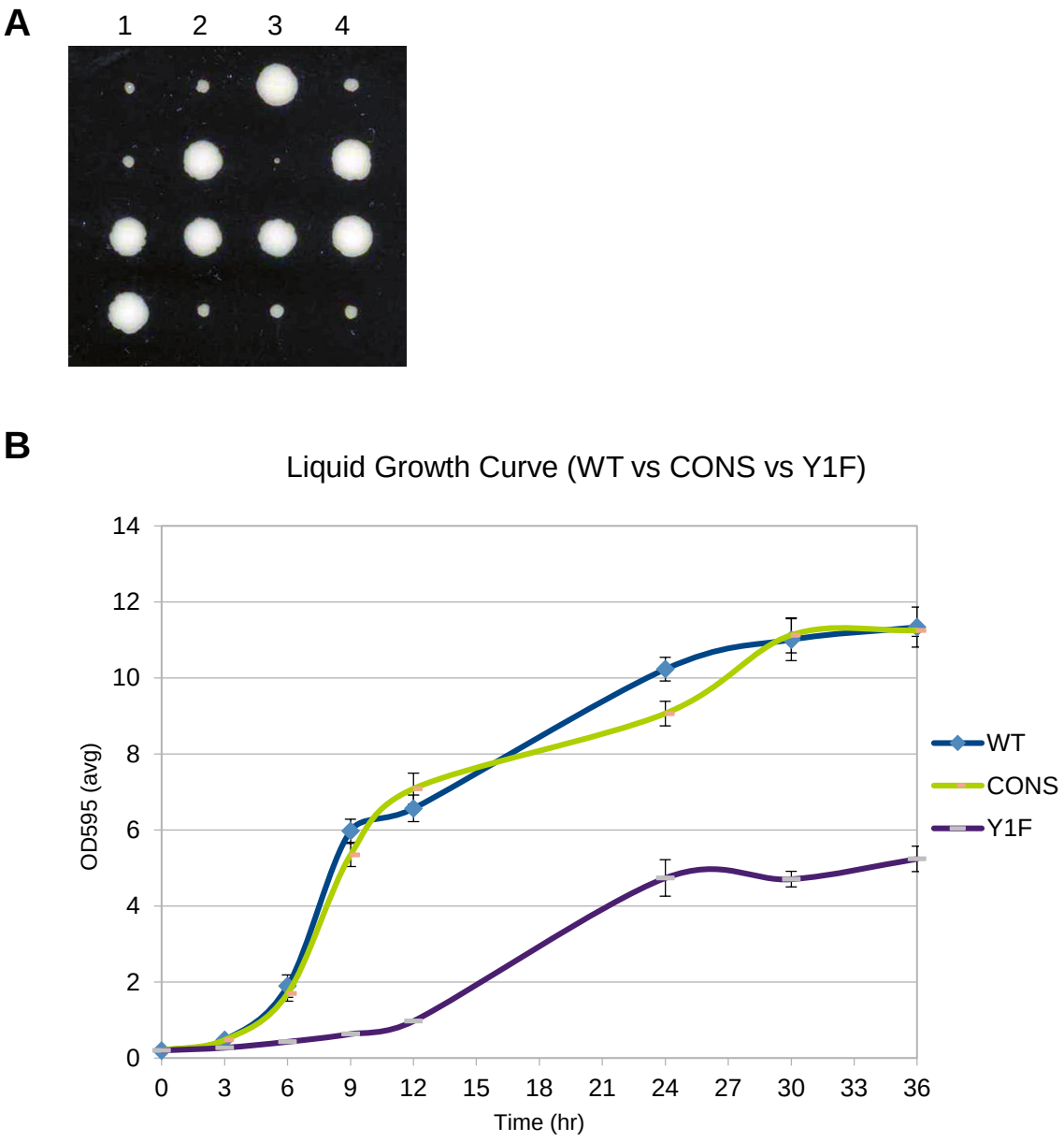


Figure S1: (A) Tetrad dissection of heterozygous diploid Y1F CTD strain (Y1F). Y1F CTD in haploid tetrads was marked with NAT resistance and was present in all small-size tetrads. (B) Liquid growth curve for WT and Y1F strains in rich medium over 48 hours. Results are averages of three biological replicates for each strain.

Figure S2

A

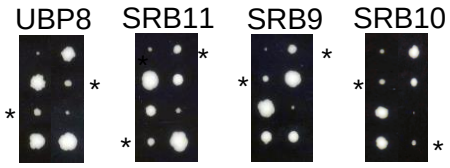


Figure S2: (A) Sample tetrad dissection results for SGA genes UBP8, SSN8, SSN2, and SSN3. Asterisk (*) indicates NAT-/G418-resistant double mutant containing both gene deletion (G418 resistance) and Y1F polymerase (NAT resistance). (B) Liquid growth curve for suppressors of Y1F strains (UBP8, SRB11, SRB9) in rich medium over 48 hours (growth of Y1F from Figure S1B is shown for comparison). Results are averages of three biological replicates for each strain.

B

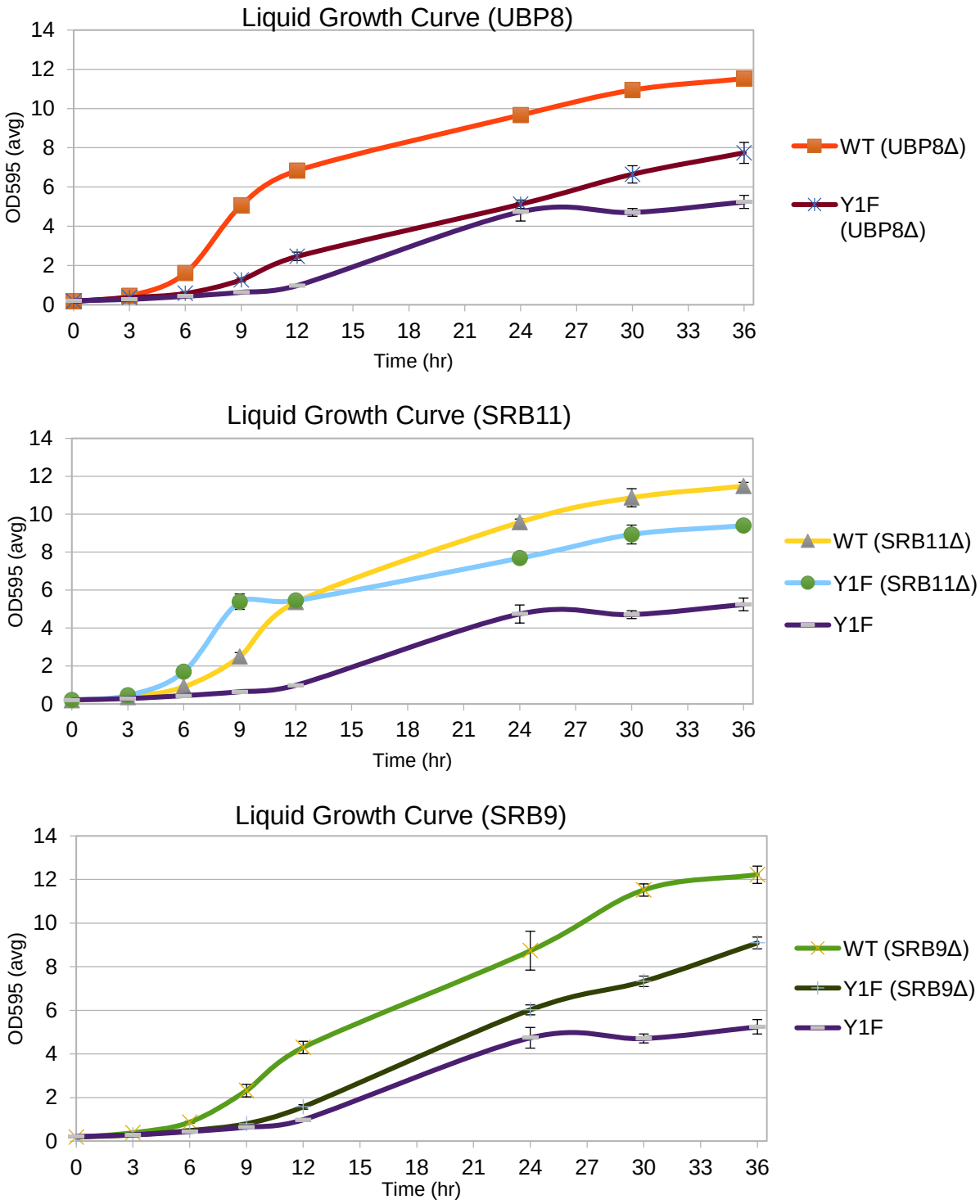
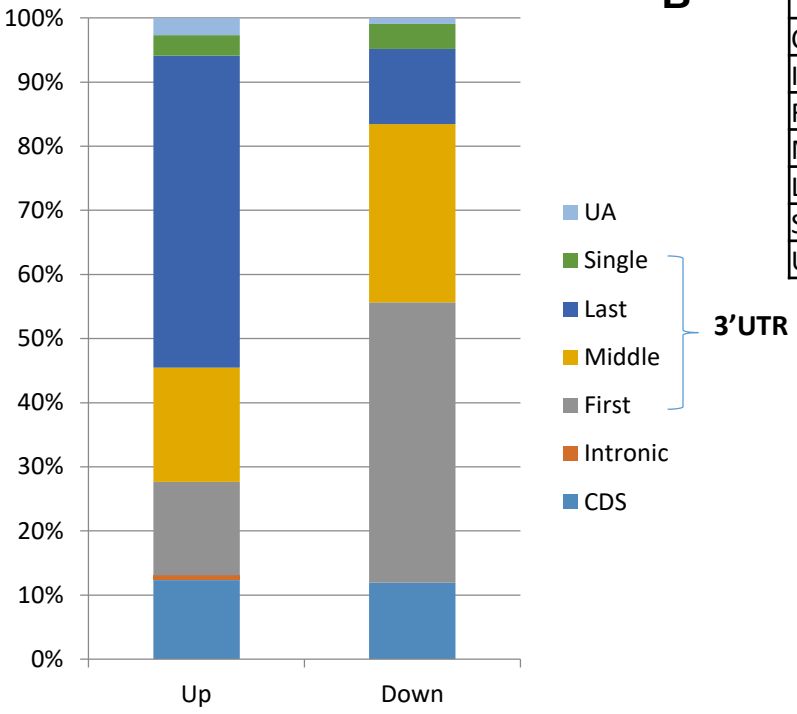


Figure S3

A



B

	Up	Down	No_sign
CDS	65	52	3052
Intronic	4	0	23
First	77	190	3089
Middle	94	121	4326
Last	257	51	3048
Single	17	17	325
UA	14	4	951

C

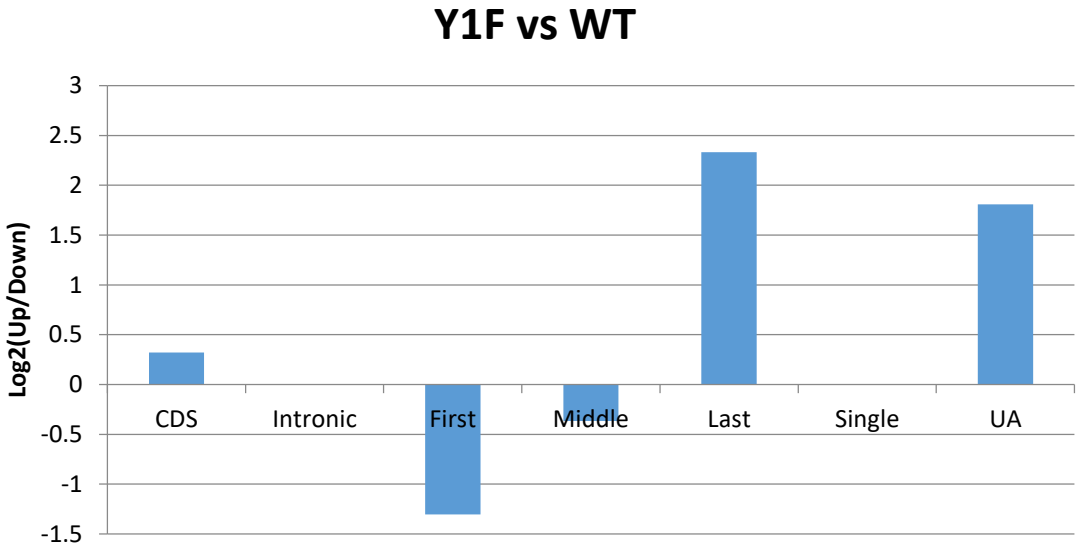


Figure S3: Additional analyses performed for 3'READS analysis. (A) Up- and down-regulation of poly(A) site usage in Y1F strain compared to WT. Labels "First" through "Single" are found in 3' UTR of genes. (B) Number of poly(A) site changes from (A) in table. (C) Log2 values for ratio of up-/down-regulated poly(A) sites in Y1F vs. WT.

Figure S4

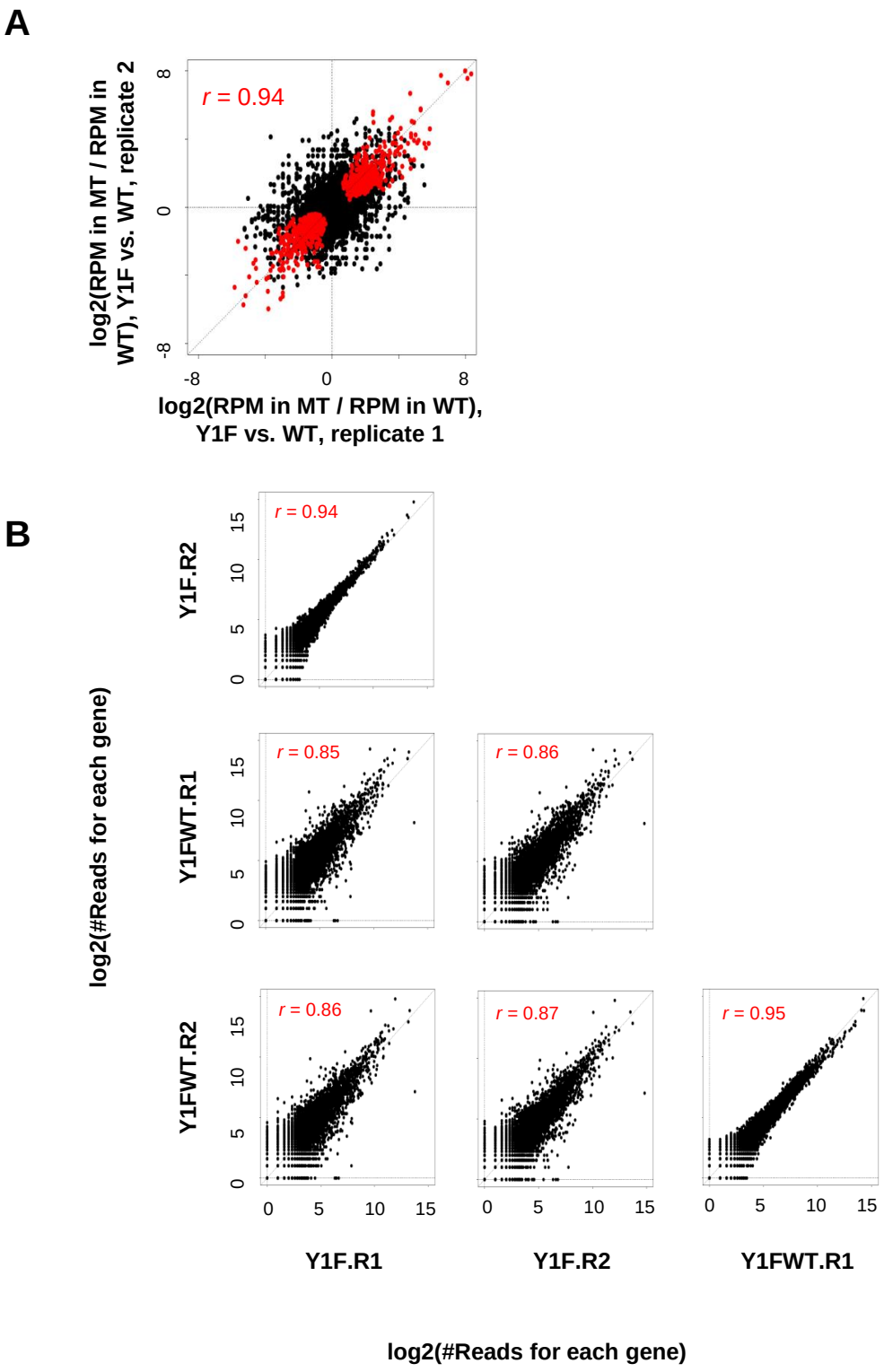


Figure S4: Correlation data for 3'READS analysis. (A) Correlation of significant genes (up-regulated/down-regulated) between WT and Y1F replicates. (B) Correlation for number of reads of genes in samples.

Figure S5

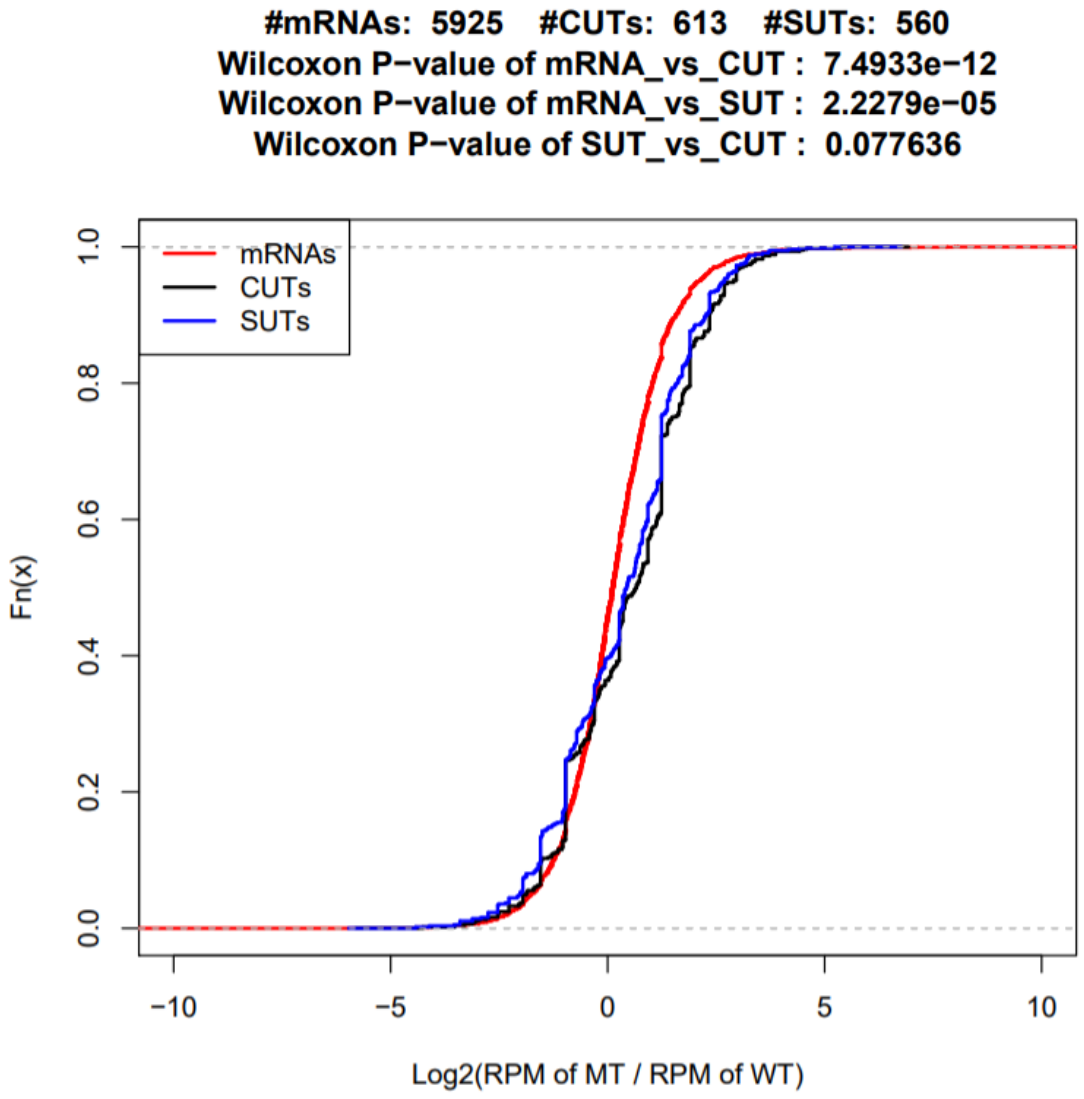
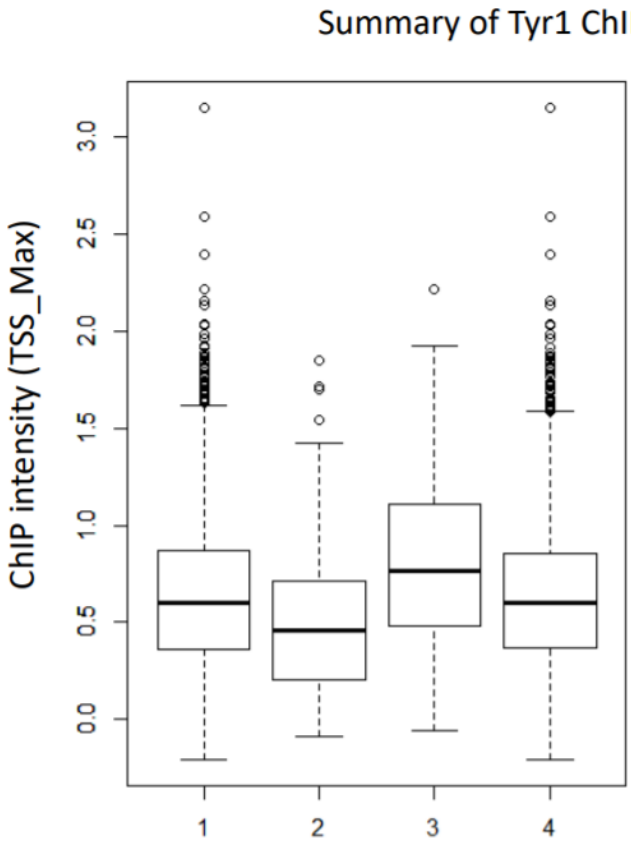


Figure S5: Analysis of CUTs and SUTs between WT and Y1F for 3'READS analysis.

Figure S6



1, All genes (4530)
2, Up-regulated genes (222)
3, Down-regulated genes (276)
4, No-changed genes (4032)

TSS_Max	Group1	Group2	Group3	Group4
Min. :	-0.2034	-0.08644	-0.05819	-0.2034
1st Qu.:	0.3649	0.20549	0.48645	0.3697
Median :	0.6005	0.45741	0.76779	0.6000
Mean :	0.6332	0.48432	0.80010	0.6300
3rd Qu.:	0.8666	0.70770	1.10623	0.8557
Max. :	3.1474	1.84801	2.21684	3.1474

1, Up-regulated gene set's Tyr1 ChIP signal is weaker than No-changed gene set.
(Welch two sample t-test, p-value = 4.449e-08)
2, Down-regulated gene set's Tyr1 ChIP signal is stronger than No-changed gene set.
(Welch two sample t-test, p-value = 6.603e-10)

Figure S6: Comparison of data between Mayer et al. 2012 and 3'READS data from this paper. Tyr1 ChIP signal intensity is compared between up-regulated, down-regulated, and genes with no expression changes in our set.

Figure S7

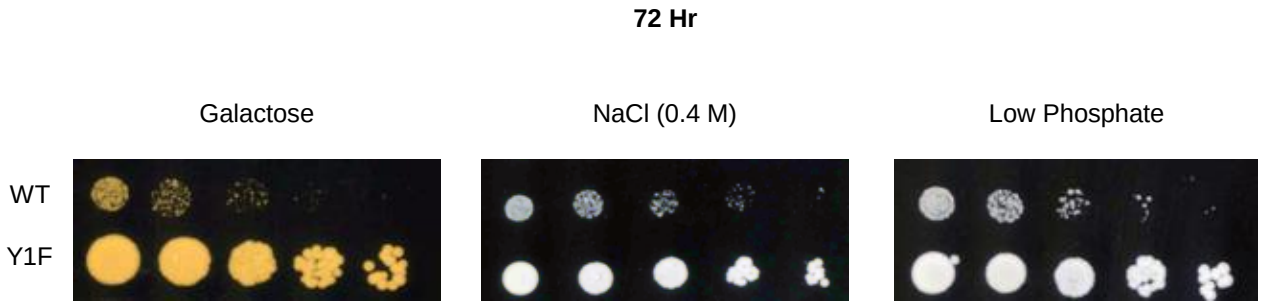


Figure S7: Additional stressors tested for changes in Y1F growth. Media containing galactose as the sole carbon source, or 0.4 M NaCl, or media depleted of phosphate were used to determine Y1F growth compared to rich media at the indicated time point. No differences from rich media were observed.

Figure S8

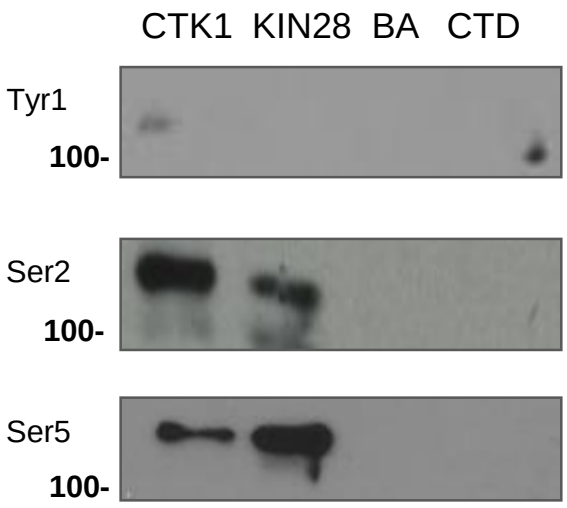


Figure S8: In vitro kinase assays with Ctk1 and Kin28. Both 3HA-tagged kinases were extracted from yeast and following IP incubated with GST-CTD as in Figure 4. Blots with Tyr1p- (3D12), Ser2p- (3E10), and Ser5p- (3E8) antibodies are shown. Lanes are Ctk1-HA (CTK1), Kin28-HA (KIN28), beads and antibody control (BA), and GST-CTD alone (CTD).

Table S1

Strain	Genotype
BY4741	MATa his3del1 leu2del0 met15del0 ura3del0
BY4742	MATalpha his3del1 leu2del0 lys2del0 ura3del0
BY4743	(Diploid) MATa/MATalpha his3del1/his3del1 leu2del0/leu2del0 lys2del0/+ met15del0/+ ura3del0/ura3del0
The following strains are haploids derived from BY4743	
NYYM203A	rpb1::rpb1-CTD(Y1F)26-NatR (Y1F 1)
NYYM204D	rpb1::rpb1-CTD(Y1F)26-NatR (Y1F 2)
NYYM205A	rpb1::rpb1-CTD(CONS)26-NatR
NYYM208A	ygl178cdel::kanMX rpb1::rpb1-CTD(Y1F)26-NatR
NYYM209A	ygl178cdel::kanMX
NYYM210C	ubp8del::kanMX rpb1::rpb1-CTD(Y1F)26-NatR
NYYM211A	ubp8del::kanMX
NYYM212B	ssn3del::kanMX rpb1::rpb1-CTD(Y1F)26-NatR
NYYM213A	ssn3del::kanMX
NYYM214A	ssn8del::kanMX rpb1::rpb1-CTD(Y1F)26-NatR
NYYM215A	ssn8del::kanMX
NYYM216B	ssn2del::kanMX rpb1::rpb1-CTD(Y1F)26-NatR
NYYM217A	ssn2del::kanMX
NYYM218D	SSN8::3HA HIS3+ rpb1::rpb1-CTD(Y1F)26-NatR
NYYM219A	SSN8::3HA HIS3+
NYYM220B	SSN3::3HA HIS3+ rpb1::rpb1-CTD(Y1F)26-NatR
NYYM221A	SSN3::3HA HIS3+
NYYM222A	SLT2::3HA HIS3+ rpb1::rpb1-CTD(Y1F)26-NatR
NYYM223A	SLT2::3HA HIS3+
NYYM224A	slt2del::kanMX
NYYM225C	slt2del::kanMX [pRS315-slt2(K54R)-3HA] LEU2+
NYYM226A	HOG1::3HA HIS3+
NYYM227A	CTK1::3HA HIS3+
NYYM228A	KIN28::3HA HIS3+
NYYM229A	hog1del::kanMX
NYYM230A	slt2del::kanMX hog1del::NatR
NYYM231C	NRD1::3HA HIS3+ rpb1::rpb1-CTD(Y1F)26-NatR
NYYM232A	NRD1::3HA HIS3+
NYYM233B	RTT103::3HA HIS3+ rpb1::rpb1-CTD(Y1F)26-NatR
NYYM234A	RTT103::3HA HIS3+
The following strains were derived from the S288C background	
Y7092	can1del::STE2pr-SpHis5 his3del1 leu2del0 ura3del0 met15del0 lyp1del LYS2+
NYYM206A	rpb1::rpb1-CTD(Y1F)26-NatR [pADH::RPB1/URA3]
NYYM207A	rpb1::rpb1-CTD(CONS)26-NatR [pADH::RPB1/URA3]

Table S1: A list of strains used for this study.